Article:

Thokala, Praveen, Ochalek, Jessica Marie orcid.org/0000-0003-0744-1178, Leech, Ashley et al. (1 more author) (2018) Cost effectiveness thresholds: the past, the present and the future. PharmacoEconomics. ISSN 1179-2027

https://doi.org/10.1007/s40273-017-0606-1

PharmacoEconomics
https://doi.org/10.1007/s40273-017-0606-1

LEADING ARTICLE

Cost-Effectiveness Thresholds: the Past, the Present and the Future

Praveen Thokala[1] · Jessica Ochalek[2] · Ashley A. Leech[3] · Thaison Tong[1]

Abstract Cost-effectiveness (CE) thresholds are being discussed more frequently and there have been many new developments in this area; however, there is a lack of understanding about what thresholds mean and their implications. This paper provides an overview of the CE threshold literature. First, the meaning of a CE threshold and the key assumptions involved (perfect divisibility, marginal increments in budget, etc.) are highlighted using a hypothetical example, and the use of historic/heuristic estimates of the threshold is noted along with their limitations. Recent endeavours to estimate the empirical value of the thresholds, both from the supply side and the demand side, are then presented. The impact on CE thresholds of future directions for the field, such as thresholds across sectors and the incorporation of multiple criteria beyond quality-adjusted life-years as a measure of 'value', are highlighted. Finally, a number of common issues and misconceptions associated with CE thresholds are addressed.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s40273-017-0606-1) contains supplementary material, which is available to authorized users.

✉ Praveen Thokala
p.thokala@sheffield.ac.uk

1 Health Economics and Decision Science, School of Health and Related Research, University of Sheffield, Regent Court, 30 Regent Street, Sheffield S1 4DA, UK

2 Centre for Health Economics, University of York, Heslington, York YO10 5DD, UK

3 Center for the Evaluation of Value and Risk in Health (CEVR), Tufts Medical Center, Boston, MA 02111, USA

Key Points for Decision Makers

This paper describes the meaning of a cost-effectiveness (CE) threshold, along with the assumptions involved using a simple hypothetical example, and highlights some of the common issues and misconceptions associated with thresholds.

CE thresholds that are being used across the world might be considered overestimates and have no empirical basis as they are based on historical estimates, heuristics or judgements.

Empirical estimates of the supply-side threshold could be considered more appropriate for judging the cost effectiveness of new technologies if the aim was to maximize population health.

1 Introduction

Cost-effectiveness analysis (CEA) is used to estimate the value for money (VfM) of new interventions in many countries across the world. In practice, the results of CEA are commonly expressed as the ratio of incremental costs to effectiveness outcomes, or incremental cost-effectiveness ratios (ICERs). Effectiveness is generally measured using a generic measure of health, typically quality-adjusted life-years (QALYs) or disability-adjusted life-years (DALYs). ICERs (i.e. cost per QALY gained or cost per DALY avoided incremental to the next best alternative) are then compared with a cost-effectiveness (CE) threshold to

Journal : Large 40273 | Dispatch : 19-1-2018 | Pages : 14
Article No. : 606 | ☐ LE | ☐ TYPESET
MS Code : PECA-D-17-00309 | ☑ CP | ☑ DISK

identify whether the new intervention is good VfM. Interventions with an ICER below a threshold (i.e. if they add each QALY or avert each DALY at a lower cost than the threshold) are considered cost effective, while those with an ICER above the threshold are not.[1]

Despite the widespread recognition and use of CE thresholds by researchers conducting economic evaluations [1, 2], as well as the adoption of their use into policy in some countries (e.g. the UK, Australia and Canada) [3–6], there is a lack of understanding among many about the meaning of thresholds, the assumptions involved, and their implications. Whether a given intervention is cost-effective or not depends on how much health it would generate and whether that amount is greater than the health that could have been generated if the money required to fund it had been spent on something else, which is a measure of opportunity cost. As such, using a CE threshold to reflect this perspective has come to be known as a 'supply side' approach [7–10]. When non-health impacts on private consumption are also considered important, some assessment of the equivalent consumption value of health is required, i.e. 'demand side' empirical research [10]. Such approaches aim to represent societal willingness to pay (WTP) for additional health gains, i.e. what individuals are willing to forego in non-healthcare/private consumption for gains in healthcare. When considering budget constraints on a healthcare system, supply-side thresholds can be considered more relevant since displacements to current health-generating interventions must happen to fund new interventions [11].

The aim of this paper is threefold: (1) to provide an illustration of the CE threshold using a hypothetical example to highlight the key assumptions involved; (2) to describe the various thresholds that are in use as policy tools in countries or have been estimated by researchers (sometimes, though not always overlapping—see the example of the UK); and (3) to present the new developments and ongoing areas of research around thresholds. The remainder of the paper is structured as follows. Section 2 presents a simple hypothetical example to illustrate how the CE threshold can be determined using the 'league table' approach, as well as optimization techniques and the assumptions involved. Section 3 describes the use of historic/heuristic estimates of the threshold, along with their limitations. Then, in Sect. 4, recent endeavours to estimate the empirical value of thresholds (e.g. work on opportunity costs in UK, Australia, Spain, as well as work in estimating thresholds for low- to middle-income countries [LMICs]) will be presented. Section 5 presents future directions for the field (thresholds across sectors such as social care, incorporating multiple criteria beyond QALYs) and their impact on CE thresholds. Finally, Sect. 6 addresses some of the common issues and misconceptions associated with CE thresholds.

2 An (Hypothetical) Example

In this section, a simple hypothetical example is used to illustrate how the CE threshold can be determined using a 'league table' approach and optimization techniques [12]. We further outline the assumptions involved in each. Let us assume there is a fixed healthcare budget of £50 million available and the aim is to choose interventions to place in the healthcare package to maximize the total QALYs gained. In this example, for the sake of simplicity, the healthcare package is empty to start with and there are seven mutually exclusive, independent interventions to choose from, each with a different set of costs and QALYs gained, as shown in Table 1. Note that these are incremental costs and QALYs associated with each intervention compared with the 'do nothing' option. At first glance, it is obvious that the budget of £50 million is not enough to fund all interventions.

2.1 League Table Approach

Under certain assumptions, which are outlined below, this 'league table' approach [13] can be used to identify the optimal allocation by including interventions according to highest VfM until the available budget is exhausted. Given our aim is to maximize health, the measure of 'value' in our example is QALYs. As we started with an empty package and are considering only independent options, we calculate VfM by dividing the costs by the QALYs of each intervention, as presented in the fourth column of Table 2 (i.e. they represent the ICERs for each intervention

Table 1 Costs and QALYs associated with the available interventions

Intervention	Cost (million £)	QALYs
A	32	7000
B	22	4000
C	20	3500
D	10	2000
E	12	1900
F	4	600
G	3	400

QALYs quality-adjusted life-years

[1] CE thresholds reflecting opportunity costs can also be used to calculate the net benefit of an intervention (i.e. if net health benefit, benefit in terms of health over and above health opportunity costs).

Journal : Large 40273 Dispatch : 19-1-2018 Pages : 14
Article No. : 606 LE TYPESET
MS Code : PECA-D-17-00309 CP DISK

Table 2 Costs and QALYs associated with the available interventions

Intervention	Cost (million £)	QALYs	Value for money (cost per QALY)	Ranking	Included in the healthcare package with a £50 million budget
A	32	7000	£4571	1	Yes (100%)
B	22	4000	£5500	3	Yes (partly funded, 36%)
C	20	3500	£5714	4	No
D	10	2000	£5000	2	Yes (100%)
E	12	1900	£6316	5	No
F	4	600	£6667	6	No
G	3	400	£7500	7	No
Maximum QALYs gained with £50 million budget: $7000 + 2000 + (0.36 \times 4000) = 10{,}454.5$ QALYs					

QALYs quality-adjusted life-years

compared with the ‘do nothing’ option). The next steps involve sorting the interventions based on their VfM and adding the interventions to the package sequentially until the budget runs out, as illustrated in the fifth and sixth columns of Table 2, respectively. This process is described in detail in the next paragraph.

As shown in Table 2, intervention A has a ranking of 1 (i.e. provides the best VfM), therefore it is added to the healthcare package first. Intervention A generates 7000 QALYs at the rate of £4571/QALY, with a total cost of £32 million, therefore there is £18 million still left from the overall budget of £50 million. The next best intervention is D, which costs £10 million and provides 2000 QALYs at the rate of £5000/QALY. After incorporating intervention D into the healthcare package, there is £8 million still left, which can be spent on the next best intervention, B. However, £8 million is not enough to fund intervention B in full (with a cost of £22 million), therefore, we can only fund a portion (8 million/£22 million = 0.36) within the budget. This would result in a gain of 1454.5 QALYs (i.e. 0.36*4000 QALYs) from intervention B at the rate of £5500/QALY. In total, we achieved 10,454.5 QALYs (7000 QALYs from A, 2000 QALYs from D, and 1454.5 QALYs from B) for the £50 million budget (see the Microsoft Excel file in the electronic supplementary material [ESM] for a visual illustration of this approach as a ‘bookshelf’) [7, 14].

In this example, the cost per QALY of the last intervention included (£5500 per QALY for intervention B) represents the supply-side threshold where that last intervention is considered ‘marginal’ (i.e. would be displaced first). The necessary assumptions required for this to be true are outlined in the ‘Underlying Assumptions’ section.

2.2 Budget-Constrained Optimization

Mathematical programming techniques can also be used to identify the optimal allocation that maximizes the total QALYs gained within the budget constraint [15, 16] (see ESM for the solution of the budget-constrained optimization problem). It can be seen that the optimal solution achieved is the same as that found using the league table approach. However, these two approaches find the same result only under a strict set of assumptions (perfect divisibility, linearity, and independence), which are described later in the ‘Underlying Assumptions’ section.

If the budget is bigger, say £51 million, we could gain a further 181.8 QALYs by spending the additional £1 million on intervention B. In fact, at the current allocation of the £50 million budget (A, B, and D), 0.0001818 additional QALYs can be gained for every £1 increase in the budget. In optimization terminology, this is termed the shadow price, i.e. how much the objective (QALYs) would increase for a one-unit increase in the constraint (budget). The shadow price can also be presented as decrements, i.e. how much the objective (QALYs) would decrease for a one-unit decrease in the constraint (budget). In our example, 0.0001818 is the shadow price of the £50 million budget optimally allocated. It should be noted that this shadow price is the inverse of the cost per QALY of the last intervention included (£5500 per QALY for intervention B). Also note that this shadow price is only applicable for a range of budget between £42 million (i.e. total costs of fully funded A and D) and £64 million (i.e. total costs of fully funded A, D, and B).

The inverse of the shadow price at the optimal allocation in the budget, referred to as the ‘critical ratio’ in one of the first mentions of the threshold in published literature [17], represents the ‘supply side’ definition of the CE threshold, i.e. a threshold representing the notion of opportunity cost. Whether a given intervention is cost effective or not thus depends on how much health it would generate and whether that amount is greater than the health that could have been generated if the money required to fund it had been spent on something else, which is a measure of opportunity cost.

Journal : Large 40273 | Dispatch : 19-1-2018 | Pages : 14
Article No. : 606 | LE | TYPESET
MS Code : PECA-D-17-00309 | CP | DISK

Table 3 Introducing intervention X to the currently optimal allocation

	Intervention	Cost (million £)	QALYs	Value for money (cost per QALY)	Ranking
Existing intervention	A	32	7000	£4571	1
Existing intervention	B	22	4000	£5500	4
Existing intervention	D	10	2000	£5000	2
New intervention	X	5.2	1000	£5200	3

QALYs quality-adjusted life-years

2.3 Assessing Cost-Effectiveness

With this allocation of the £50 million budget, for a new intervention X to be included in the healthcare package (X does not have to be from the existing list in Table 1), we would need to disinvest first (assuming the overall budget is fixed at £50 million). This disinvestment is only worthwhile if the replacement of existing interventions in the current healthcare package with X brings positive net QALYs gained. Let us introduce a new intervention X that costs £5.2 million and provides 1000 QALYs at the rate of £5200 per QALY (Table 3). Given the existing allocation (A, B and D), the decision is whether we should fund X.

As illustrated earlier, the threshold is the inverse of the shadow price of the budget with its current optimal allocation, which is £5500 per QALY (i.e. the cost per QALY of the last intervention included, intervention B). Since £5200/QALY (VfM of X) is lower than £5500/QALY (current inverse of the shadow price), it is cost effective to replace B with X. That is, more QALYs can be gained by spending money on X than those lost by displacing part of B. In this case, X will be funded from the replacement of part of B, the proportion of B left after funding X is estimated as follows: (£8–5.2 million)/£22 million = 12.7%. Replacing B with X would generate 1509.1 QALYs for the £8 million (i.e. 1000 QALYs from X + 0.127*4000 QALYs from part of B). In total, we achieve 10,509.1 QALYs (7000 QALYs from A, 2000 QALYs from D, 1000 QALYs from X, and 509.1 QALYs from B) for the £50 million budget, an increase of 54.6 QALYs (10,509.1 − 10,454.5 = 54.6 QALYs) compared with the previous allocation.

2.4 Underlying Assumptions

Through this example, we illustrate below a few key assumptions relating to CE thresholds that are worth further consideration, i.e. perfect divisibility, linearity, independence, marginal increments in budget, disinvestment plan, perfect information and other issues [14, 18].

2.4.1 *Perfect Divisibility, Linearity and Independence*

One assumption that applies to both the league table approach and the budget-constrained optimization example is the notion of perfect divisibility (i.e. a proportion of the intervention can be funded if there are not enough funds to cover the costs of the whole intervention). In the above example for the optimal allocation (before X was introduced), the £8 million left was not enough to cover the whole of intervention B (£22 million) and it was assumed that intervention B can be funded in part (0.36 = £8 million/£22 million) within the remaining budget, resulting in a gain of 1454.5 QALYs from B (assuming linearity, i.e. increase in costs results in a proportional linear increase in QALYs, also known as ‘constant returns to scale’). It should be noted that the assumption of perfect divisibility may not always hold true in real life; for example, if there is a need for expensive specialist equipment, it must be purchased in full as a fraction of equipment cannot be bought. In addition, while the perfect divisibility may be achieved by limiting the patient population receiving the technology (e.g. by subgroup), the linearity assumption may not be valid (e.g. as the costs and QALYs for the subgroup may be different from the overall population).

It should be noted that the league table approach cannot be used if the perfect divisibility assumption does not hold. In case of the optimization, the problem needs to be solved again using integer constraints. In the above example, the resulting optimal solution with integer programming (before X was introduced) is to fund interventions A, D, F and G in full to achieve 10,000 QALYs for a budget of £49 million (see the Integer Optimization sheet in the Microsoft Excel file in the ESM). This is because even though there are interventions with better VfM than F and G, they are not affordable within the leftover available budget after funding A and D (i.e. interventions B, C and E cost more than £8 million).

Similar issues arise when considering interventions that are interdependent—VfM techniques are not applicable and optimization techniques should be used to account for the interactions [19]. These issues arise because the league table approach assumes perfect divisibility, linearity and independence and is based on the use of cost per QALY ratios without considering budget impact. While the optimization problem can be structured using integer programming to overcome these issues, the shadow prices are

Journal : Large 40273 | Dispatch : 19-1-2018 | Pages : 14
Article No. : 606 | ☐ LE | ☐ TYPESET
MS Code : PECA-D-17-00309 | ☑ CP | ☑ DISK

no longer applicable for these methods (and thus the thresholds are not easily interpretable).

2.4.2 *Marginal Budget Impact*

The threshold, the inverse of the shadow price or the cost per QALY of the last intervention included, is only applicable for interventions with a small impact on budget, typically termed 'marginal' impacts on budget. In the example above, the new intervention X had a budget impact of £5.2 million, which meant only intervention B needed to be displaced, hence the threshold of £5500/QALY. If the budget impact of X was high (which in our example is any amount above £8 million, the money spent on intervention B), it would be necessary to consider whether it is cost effective to also replace the next existing intervention in the package (intervention D) with X since there is still room to fund more X. Now, the £5500 per QALY from the inverse of the shadow price is no longer applicable.[2] We need to compare the VfM of X (£5200 per QALY) with that of D (£5000 per QALY). Since £5200 per QALY is greater than £5000 per QALY, X should not replace D. Thus, as seen in the above example, while the threshold can be considered appropriate at marginal impacts on budget, the value of the threshold needs to be more conservative for interventions with higher budget impacts to accommodate the displacement of more cost-effective interventions. As such, many countries have started to impose a 'budget impact limit' alongside CE considerations (see Sect. 6.4).

2.4.3 *Disinvestment Plan*

In our example, we assume that the disinvestment to fund a new intervention should come from the least cost-effective intervention(s). The new intervention was only compared with the least cost effective existing intervention within the optimal allocation, to keep with our original aim of maximizing QALYs. Replacing interventions other than the least cost-effective intervention (i.e. anything other than the intervention with least VfM) in our healthcare package will result in greater QALYs lost than when displacing least cost-effective intervention. However, it is not always possible to ensure that the least cost-effective intervention(s) are disinvested first or that the healthcare package is 'optimal' [20]. Healthcare packages in real-life settings tend to include a mix of interventions that are cost effective as well as cost ineffective, and there might not be information on what interventions are being displaced. Thus, the empirical estimates of the 'supply-side threshold' use marginal productivity of the system, which describes the relationship between changes in healthcare expenditure and health outcomes (i.e. change in the QALYs of the healthcare system with change in the budget—see Sect. 4.1).

2.4.4 *Perfect Information (and Other Assumptions)*

In our example, we assume that we start with an empty healthcare package and that the information (i.e. the overall budget, the interventions available, and the data on costs and QALYs for all interventions) is already known. Our example is a very simple approximation, whereas the reality of healthcare resource allocation is much more complex. For instance, the budget may vary with time (and in fact there could be different budgets to consider); there may be complementarities between interventions (e.g. early diagnostic interventions would improve the benefits of treatment interventions, violating the independence assumption); and the healthcare package may already include many pre-existing interventions (where the implications of disinvestment may need to be considered first). Furthermore, full knowledge of costs and benefits for all interventions required to estimate the threshold value is usually incomplete (i.e. the data required, either to develop the comprehensive league table or to formulate the optimization problem, to determine the threshold value is not available).

3 Past: Use of Heuristics/Historical Estimates of Thresholds

Given the challenges highlighted in specifying a threshold consistent with QALY maximization in the earlier section, many countries use a threshold value based on other methods and representing different concepts. For example, in line with previous WHO-CHOICE guidance [8, 21], some LMICs have employed a heuristic of one to three times the gross domestic product (GDP) per capita [22, 23], while the UK, Ireland and the US use explicit thresholds broadly based on historical estimates/judgement [24, 25]. Many countries (including Canada, Brazil, Australia, and Sweden) do not specify an explicit threshold at all [4, 26]. This section briefly summarises how the thresholds based on heuristics or historical estimates, whether explicit or implied, are used across the world.

[2] As described earlier, the shadow price of 0.0001818 relates to intervention B and as such is only applicable for a range of budgets between £42 million (i.e. total costs of fully funded A and D) and £64 million (i.e. total costs of fully funded A, D and B).

Journal : Large 40273 Dispatch : 19-1-2018 Pages : 14
Article No. : 606 ☐ LE ☐ TYPESET
MS Code : PECA-D-17-00309 ☑ CP ☑ DISK

3.1 Explicit Thresholds

3.1.1 UK (National Institute for Health and Care Excellence)

The National Institute for Health and Care Excellence (NICE) in the UK is a high-profile example of the use of explicit CE thresholds, and its guidance recommends in favour of funding interventions with an ICER below a threshold of £20,000/QALY or £30,000/QALY, and also recommends against funding interventions with an ICER above these thresholds [27–29]. However, a higher threshold (i.e. £50,000/QALY) is used for life-extending treatments for small patient populations at the end of life, i.e. treatments that offer an extension to life greater than 3 months compared with current treatment in the National Health Service (NHS); are for patients with a short life expectancy, i.e. normally < 24 months; and are for small patient populations, normally not exceeding a cumulative total of 7000 patients for all licensed indications in England [30]. Despite this guidance, interventions with ICERs above £30,000 or £50,000 are often accepted, even when lacking the requisite special evidence needed [31].

3.1.2 Ireland

The CE of all new medicines in Ireland is considered by the National Centre for Pharmacoeconomics (NCPE), in collaboration with the Health Service Executive (HSE), the public body with responsibility for delivering state-funded healthcare in Ireland. The Irish Pharmaceutical Healthcare Association (IPHA) and HSE have an agreement that explicitly states that the QALY threshold to be used in the HTA process is €45,000 [32]. This value is also confirmed on the NCPE website [33]. It is worth noting that, unlike NICE, NCPE's recommendations are not mandatory and can be overruled by the minister/HSE [25].

3.1.3 US

While \$50,000 per QALY has been mentioned anecdotally in the past in the US [34], the recent value frameworks mention explicit thresholds. Given the diversity of payers and healthcare organizations, it should be noted that there are differences in the thresholds used. A high-profile example of explicit reference to thresholds is the use of \$100,000–\$150,000/QALY for a value-based price benchmark by the Institute for Clinical and Economic Review (ICER) [35], a trusted non-profit organization that evaluates evidence on new technologies in the US. Premera Blue Cross, a large not-for-profit health plan in the Pacific Northwest, uses value-based formulary tiers based on ICER thresholds—drugs are allocated to one of the four co-payment tiers (tier 1, < \$10,000/QALY; tier 2, \$10,000 to < \$50,000/QALY; tier 3, \$50,000 to < \$150,000/QALY; and tier 4, > \$150,000/QALY) [36].

3.2 Heuristics for the Threshold Value: WHO-CHOICE (One to Three Times a Country's Gross Domestic Product)

One to three times a country's annual GDP per capita has been a widely used threshold for CE studies within global health, mainly among studies focused on LMICs [1, 37]. A recent study found that the proportion of LMICs citing this threshold has substantially increased over time, with 10% of studies citing this threshold in the early 2000s, to 76% between 2013 and 2015 [37]. While the origins for its intended use for CEA are less clear, the WHO first used these values in its 2001 Commission on Macroeconomics and Health (CMH) report [38]. While this report intuitively equates a year of life to per capita income, considering productivity and leisure time, it used per capita income to value the economic loss resulting from the burden of major diseases impacting countries. Despite its variant aim, the WHO-CHOICE thereafter adopted this range for promoting CEA [21, 38]. There have recently been several opinions on this threshold value that have motivated calls for consensus and new primary research [8, 9, 37, 39–43]. For instance, some analysts have argued that CE thresholds reflecting opportunity costs are much lower than the one to three times GDP per capita rule of thumb, while other analysts encourage applying a range of income elasticity estimates to account for the relationship between the value per statistical life (VSL) and income [39]. The WHO has since backed away from this threshold range and recognizes its limitations for CEA [8].

3.3 Implied/Unspecified Thresholds

A recent systematic overview of CE thresholds suggested that many countries do not specify a threshold [26]. While researchers analysed previous decisions to identify the threshold value in these countries, they were unable to pin down a single number. Nevertheless, the manner in which these countries use different CE thresholds is briefly described below.

3.3.1 Pharmaceutical Benefits Advisory Committee

The Pharmaceutical Benefits Advisory Committee (PBAC) in Australia does not formally specify a CE threshold. However, the cost per QALY of the technology is reported as belonging to one of four bands, i.e. AUS\$15,000–\$45,000; \$45,000–\$75,000; \$75,000–\$105,000; \$105,000–\$200,000. A recent study by Paris and Belloni [44] at the

Journal : Large 40273 | Dispatch : 19-1-2018 | Pages : 14
Article No. : 606 | LE | TYPESET
MS Code : PECA-D-17-00309 | CP | DISK

Organisation for Economic Co-operation and Development (OECD) suggested that technologies with ICERs greater than $75,000/QALY were rarely recommended and those greater than $45,000/QALY were recommended only in exceptional circumstances, where there was high clinical need and no alternative treatment. These findings are similar to those observed by Henry et al. in their retrospective analysis of PBAC decisions [45].

3.3.2 Canadian Agency for Drugs and Technologies

While the Canadian Agency for Drugs and Technologies in Health (CADTH) guidelines for the economic evaluation of health technologies recommend the use of a 'supply-side' estimate of the CE threshold, that value is not given in the guidance [4]. While the reporting sometimes refers to the $50,000/QALY threshold (for example, the probability of being CE was *x*% at a threshold of $50,000/QALY), a review of all the publically available CADTH appraisals performed by Griffiths and Vadlamudi [46] suggested that this threshold is not consistently applied, with several technologies recommended with ICERs above $50,000 per QALY, while many were rejected with ICERs below this threshold.

3.3.3 New Zealand

The Pharmaceutical Management Agency (PHARMAC) in New Zealand state that they do not have a CE threshold [47]. While researchers have tried to imply the threshold from previous decisions [48, 49], PHARMAC states that they fund medicines within a fixed budget, and as CE is only one of its nine decision criteria used to inform decisions, thresholds cannot be inferred or calculated [50]. They also note that CE estimates for PHARMAC's investments has ranged between −NZ$40,000 (net cost savings to the health sector for health gains) to over +NZ$200,000 per QALY (−€20,000 to +€100,000) [51].

3.3.4 Other Countries

Other countries, including Scotland [52], Korea [53] and Brazil [54], use CE analyses for decision making but do not explicitly specify a threshold.

4 Present: Empirical Estimates of Cost-Effectiveness (CE) Threshold

Recently, some countries have begun to conduct empirical research to identify CE thresholds for their setting. These studies have broadly been classified as either supply- or demand-side estimates [10]. Supply-side estimates aim to reflect the opportunity cost of spending on health by linking the healthcare expenditure to health outcomes, while the demand-side estimates aim to reflect societal WTP for improvements in health.

4.1 Supply-Side Thresholds

It should be noted that the example in Sect. 2 illustrates an ideal situation in which the budget allocation is optimal; it is easy to identify the least cost-effective intervention(s), and the system (decision makers) only displace these least cost-effective interventions. This is a 'first best' situation; however, in practice, this is not always the case. In complex systems, the existing healthcare package may not be optimal, it may not be possible to specify exactly what activities are displaced, and decisions about disinvestment may be left to other decision makers in the system, for example at a local level. Thus, in empirically estimating the threshold, the aim is to estimate the shadow price of the budget in terms of the interventions that are likely to be displaced [42]. This is what Culyer describes as an approach to estimating the 'second best' threshold [7]. These empirical estimates of the supply-side threshold tend to reflect the marginal productivity of the healthcare system, derived from the relationship between changes in healthcare expenditure and health outcomes, where expenditures at the margin may be committed to a mix of cost-effective and cost-ineffective interventions (i.e. interventions with a range of cost per QALYs) [7]. In a world where the assumptions of the optimization model are met, this conceptualization of the threshold should result in the same value as that which is derived by solving the constrained optimization problem. However, where the necessary assumptions as set out in the preceding section are not met, the values may differ. The 'second best' approach provides an estimate that best informs the expected health opportunity costs of a new intervention and, therefore, if robustly estimated, can be better relied on to inform whether a new intervention is expected to result in a net health gain or net health loss.

There are challenges involved in estimating the relationship between changes in healthcare expenditure and health outcomes, i.e. the marginal productivity of the healthcare system. Given the outcome of interest is QALYs [a combination of quality of life (QoL) and life-years (LYs)], there is a need to link the healthcare expenditure to mortality (to estimate the effect on LYs) and morbidity (to estimate the effect on QoL). The data on healthcare expenditure and its effect on mortality/morbidity may not always be readily available and, as such, assumptions are often required. Furthermore, there are also econometric challenges that include, but are not limited to, issues around controlling for the many non-healthcare factors that

Journal : Large 40273 | Dispatch : 19-1-2018 | Pages : 14
Article No. : 606 | ☐ LE | ☐ TYPESET
MS Code : PECA-D-17-00309 | ☑ CP | ☑ DISK

affect health [55], which if not properly accounted for may lead to biased and inconsistent estimates. To date, such within-country estimation has been undertaken in relatively few countries, which are described below alongside one example where cross-country data have been used to estimate these values for a number of countries.

4.1.1 UK

Claxton et al. [42] empirically estimated the CE threshold for the NHS in the UK to be £12,936 per QALY. They used the English NHS programme budgeting data to estimate the relationship between changes in overall NHS expenditure and changes in mortality/LYs gained, and subsequently extended this to QALYs. Their ‘structural’ uncertainty suggested that the estimate is likely to be an overestimate and reported that the probability the threshold is less than £20,000 per QALY is 0.89, and the probability that it is less than £30,000 per QALY is 0.97. The assumptions made in the estimation of the UK threshold have been discussed in a number of publications [28, 56, 57].

4.1.2 Australia

Edney et al. [58] estimated the CE threshold, called the reference ICER, for Australia. They used an instrumental variable two-stage least squares regression to estimate the effect of changes in health expenditure on QALYs due to reduced mortality. Further empirical analysis was then used to inform the effect of health expenditure in terms of QALYs due to reduced morbidity. These are then combined to produce a central estimate of the reference ICER, which represents the average opportunity costs of decisions to fund new technologies, i.e. AUS$28,033/QALY.

4.1.3 Spain

Vallejo-Torres et al. [59] estimated the CE threshold for the Spanish NHS. They used 5 years of data across the 17 regional health services in Spain to regress quality-adjusted life expectancy (QALE) against health spending, controlling for region and year fixed effects, and a comprehensive set of time- and region-variant indicators, applying a 1-year lag to expenditure. They report that health expenditure has a positive and significant effect on QALE, with an average spending elasticity of 0.07, which translates into a cost per QALY of between €21,000 and €24,000.

4.1.4 Low- to Middle-Income Countries CE Thresholds

Ochalek [41] estimate CE thresholds for 123 LMICs using estimates of the effect of a change in government spending on health on health outcomes from cross-country data. Their study expands on existing studies within the literature estimating the effect of a change in spending on mortality outcomes to estimate the effect of a change in spending on a range of mortality and morbidity outcomes. Using data on each country’s demography (i.e. the sex and age structure of the population), epidemiology (i.e. underlying mortality and morbidity burden) and health expenditure, they were able to generate a range of cost per DALY averted estimates for 123 countries that captures some of the structural uncertainty associated with these estimates. Their results aim to reflect the rate at which the healthcare system in a given country is able to produce health, and, as such, can be used to inform health opportunity costs. For example, they have been used to help guide decisions around the design of the Essential Health Package in Malawi [60, 61].

4.2 Demand-Side Thresholds

The empirical methods of estimating demand-side thresholds, namely WTP and value of a statistical life studies, are reviewed and discussed in detail by Vallejo-Torres et al. [10]. Below, we offer a brief description of the application of these methods in policy in two countries—Thailand and Malaysia.

4.2.1 Thailand (Health Intervention and Technology Assessment Program)

The Health Intervention and Technology Assessment Program (HITAP) in Thailand elicited the WTP for a QALY in the Thai healthcare setting [62]. The results of this study were adopted by decision-making bodies as the appropriate threshold for health investment in the Thai setting; the ceiling threshold is reported to be 160,000 Baht per QALY, which is approximately 1.2 times Gross National Income (GNI) per capita [63]. However, they also note that this single threshold is not used for resource allocation of all types of interventions; for example, sometimes medicines that treat rare diseases are included in the National List of Essential Medicines (NLEM) even though their ICER is much higher than the threshold.

4.2.2 Malaysia

Lim et al. conducted a cross-sectional, contingent valuation study in four states of Malaysia to estimate the CE threshold for healthcare interventions as WTP for a QALY [64]. One thousand and thirteen respondents were interviewed in person for their socioeconomic background, QoL, and WTP for a hypothetical scenario. The authors reported that the CE thresholds ranged from MYR12,810 to MYR28,470 (US$4000–US$8900) and education level,

Journal : Large 40273 | Dispatch : 19-1-2018 | Pages : 14
Article No. : 606 | ☐ LE | ☐ TYPESET
MS Code : PECA-D-17-00309 | ☑ CP | ☑ DISK

estimated monthly household income, and the description of health state scenarios had the biggest effect on the WTP estimates. They concluded that there is no single WTP value for a QALY and that the CE threshold estimated for Malaysia was found to be lower than the threshold value recommended by the WHO (i.e. one and three times the GDP per capita, which was approximately $10,000 and $30,000, respectively, in 2017) [65].

5 Future: Beyond Quality-Adjusted Life-Years (QALYs)? Other Sectors?

Most of the work on CE thresholds has been based on using QALYs (or DALYs) as the measure of effectiveness. However, there have recently been some developments that suggest an inclination to go beyond these measures of health benefit, including the recent work on value frameworks [66], which mentions a number of additional criteria in addition to QALYs or DALYs, and the recommendation statement from the Second Panel on Cost-Effectiveness in Health and Medicine [67, 68], which supports the use of a societal perspective. The impact of these recommendations is discussed in brief below.

5.1 Thresholds for Benefits Beyond QALYs

Alongside the recent work on value frameworks [66], which mentions many additional criteria beyond QALYs, it is widely acknowledged that many HTA organizations consider multiple factors alongside CE [69]. More recently, there have been calls for including these multiple criteria explicitly in the assessment of value [70], using techniques such as multicriteria decision analysis (MCDA) [71]. The current CE thresholds are based on QALYs (or DALYs) being the measure of effectiveness. If the value is redefined to include multiple criteria beyond QALYs (or DALYs), the measure of effectiveness is not QALYs (or DALYs) anymore but rather a new composite measure of effectiveness. As such, the threshold will need to be re-estimated for this new measure of 'effectiveness' to reflect the opportunity costs [6]. As observed in Sect. 4 (the empirical estimates of the supply-side thresholds), this poses a significant informational challenge in identifying the marginal impacts on the different criteria that make up the overall effectiveness.

5.2 Thresholds in Other Sectors

The Second Panel on Cost-Effectiveness in Health and Medicine [67, 68] supports a societal perspective and recommends the use of an 'impact inventory'—a structured table listing the health and non-health effects of an intervention that should be considered in a societal reference-case analysis. To evaluate interventions crossing multiple sectors, sector-specific thresholds are needed that represent the sector-specific outcome that would be forgone as the result of the additional costs of a new intervention. To date, no sector outside of healthcare has established a threshold. While some sectors have established measures, such as the Adult Social Care Outcomes Toolkit (ASCOT) used to estimate social care-related QoL (SCRQoL), many sectors do not have standard definitions for their outcomes. The challenges involved in performing CEA when the intervention concerns multiple sectors are highlighted by Remme et al. [72].

6 Key Issues/Misconceptions with Thresholds

6.1 Which Thresholds Should be Used?

Unless there is clear reason to choose a different threshold value (e.g. political sensitivity), empirical estimates provide a more appropriate value of the threshold than historical/heuristic thresholds, which are based on judgement. The key question is whether supply-side thresholds (which aim to represent the opportunity cost of investment to the system, given budget constraints) or demand-side thresholds (WTP estimates that aim to reflect the value that society places on a QALY) should be used [10]. A recent systematic review of WTP per QALY studies suggested that WTP per QALY varied substantially by condition, especially those for extending or saving life and improving QoL [73]. Supply-side thresholds enable the quantification of the net health gains (or losses) that would result from the inclusion of a new intervention (whether doing so represents an increase in the budget or displaces a currently funded intervention[s]) in the healthcare system. Decisions made on the basis of supply-side CE thresholds ensure that aggregate health is improved by the inclusion of new interventions.

On the other hand, thresholds based on WTP for a QALY are generally higher than thresholds resulting from estimating the opportunity cost to the healthcare system [10]. As such, using WTP estimates may lead to decisions that reduce rather than improve health outcomes overall. This may also be the case with the use of WHO-CHOICE guidelines for thresholds (i.e. one to three times the GDP), where the threshold is not related to the efficiency of the healthcare system. However, as WTP estimates reflect societal WTP for improvements in health, the fact that they tend to be higher than estimates linked to the efficiency of the healthcare system provides suggestive evidence for an increase in public budgets for healthcare. Some analysts have argued that in a privately funded healthcare system, in

Journal : Large 40273 | Dispatch : 19-1-2018 | Pages : 14
Article No. : 606 | LE | TYPESET
MS Code : PECA-D-17-00309 | CP | DISK

the absence of explicit healthcare budget constraint, WTP can be an estimate of the opportunity cost of private consumption [2].

6.2 Should the Threshold be Made Explicit?

There are two questions here: (1) whether there can be a single threshold, and (2) whether the threshold values should be made public. No HTA organization currently recommends the use of a single threshold, and many do not explicitly specify a threshold at all (as seen in Sect. 3). Those that specify a threshold tend to specify a range rather than a single value reflecting the belief that a single threshold should not be applied to the diverse range of technologies and conditions. In terms of the second question, the so-called 'silence of the lambda' [74] or reluctance to set out an explicit threshold, may result from a number of concerns, including fear of gaming by pharmaceutical companies to target ICERs just below the threshold, reduced flexibility to balance competing criteria when making funding decisions, and the issues associated with advocating a threshold value that may have little or no empirical basis (such as the potential for political and ethical concerns about the accuracy and validity of funding decisions) [75].

6.3 Impact of Using the Wrong Threshold

If the threshold used is lower than the empirical estimate, it may lead to potentially cost-effective (compared with the empirical threshold) technologies not being reimbursed. However, it should be noted that in situations where researchers suggested increasing the threshold [76], arguments were based on WTP/preference estimates. On the other hand, if the threshold used is higher than the empirical estimate reflecting health opportunity costs, each new technology approved (with a higher ICER than the empirical threshold) leads to loss in health outcomes. An example is NICE's end-of-life decision-making scheme, where it was suggested that approving drugs with an ICER higher than the NICE threshold of £20 000–£30 000/QALY resulted in substantial QALY losses [77]. Furthermore, Claxton et al. [31] argue that the current NICE threshold (of £20 000–£30 000/QALY) is too high compared with the empirical estimates, suggesting that approving drugs lead to more health likely to be lost than gained.

6.4 Threshold and Budget Impact

If the budget impact of a new technology is substantial (i.e. non-marginal), the threshold used should be lower, reflecting the size of the budget impact, as the new technology will displace a large proportion of the existing health services (see example in Sect. 2.1) [78]. The recent hepatitis C drugs highlight this issue; while the new hepatitis C drugs were very cost effective, their budget impact was quite substantial [79]. ICER in the US has a limit for budget impact ($915 million/year for 2017–2018) designed to alert policy makers that funding the new service may be difficult without displacing other needed services or increasing the healthcare insurance costs [35]. In the UK, for cost-effective technologies with significant budget impact (NICE use a 'budget impact threshold' of £20 million per year), special arrangements need to be agreed in dialogue with companies to better manage the introduction of these technologies in the NHS [80].

6.5 Threshold and Inflation

Many have argued for a higher threshold as the values used by NICE, PBAC, the US, etc., have remained the same since they were first introduced [81]. In the absence of an explicit healthcare budget constraint, inflation can potentially affect the WTP estimates of the threshold; however, if the threshold is linked to the efficiency of the healthcare system (i.e. CE of the displaced services), it is not related to inflation. If a health service became more efficient over time (i.e. the displaced activities become more cost effective over time), the threshold will fall irrespective of inflation. This argument is also applicable for the transferability of thresholds between countries. Rather than relying on generic metrics such as GDP (e.g. WHO-CHOICE guidelines for thresholds of one to three times the GDP) or exchange rates, the thresholds should be determined by estimating the efficiency of the healthcare system, as observed in Sect. 4.

6.6 Threshold and Capacity Constraints

Published CEA studies often ignore the capacity constraints of resources (e.g. beds, nurses, equipment, etc.), which may result in biased estimates of CE [82]. In principle, if perfect information was available, these capacity constraints can be added, on top of the budget constraint, into the optimization problem to estimate the 'new' CE threshold that takes into consideration the scarcity of resources. However, this perfect information is not available in reality and thus these capacity constraints are incorporated within CE modelling to understand their impact on the standard of care and the implementation of the new technology [83].

Where perfect information about capacity constraints does not exist, empirically estimated 'supply side' CE thresholds can be used to determine the expected value of reducing or removing such constraints, either specific to

Journal : Large 40273 | Dispatch : 19-1-2018 | Pages : 14
Article No. : 606 | LE | TYPESET
MS Code : PECA-D-17-00309 | CP | DISK

interventions or across the healthcare system as a whole. This expected value can be used alongside information (e.g. based on expert opinion) about the costs and benefits of removing different constraints to prioritise policies to reduce or remove constraints to scale up the implementation of interventions [61].

6.7 Priority-Setting Process

Alongside the results of CEA, a number of other factors are often also considered as part of the appraisal process around whether to adopt or reject an intervention. A recent review of all HTA appraisals between May 2000 to May 2014 from the NICE, PBAC, Scottish Medicines Consortium, and CADTH suggested that technologies with ICERs higher than the respective thresholds are sometimes recommended; the reasons included high clinical benefit over the standard of care, and addressing an unmet therapeutic need [84]. Similarly, even though some technologies (such as orphan drugs for rare diseases, or cancer treatments at end of life) have very high ICERs, NICE and most other health systems have found ways to fund those few technologies on the basis of evidence of benefit. On the other hand, some interventions are rejected, even when the ICERs are below the threshold [46]. Indeed, it is acknowledged that there is a need for some discretion in priority setting linked to legitimation of decisions rather than using the threshold alone.

7 Conclusions

This paper contributes to the literature on CE thresholds by providing a simple illustration of the CE threshold as the shadow price of budget constraint, providing a theoretical framework for how a CE threshold could be employed in a hypothetical optimization setting. Existing estimates of 'thresholds' representing various definitions, from heuristics applied historically to more recent empirical estimates, whether WTP for improvements in health or opportunity costs are then outlined. Among these, those that can be categorized as supply-side estimates (i.e. from the UK, Australia, Spain and LMICs, as presented in Sect. 4.1) may be considered more appropriate for judging the CE of new technologies where the aim of agencies is to inform whether or not a new technology is expected to improve population health. Finally, the future for CE thresholds is speculated upon where new policy questions have indicated further areas of research where thresholds will be relevant and useful for decision making, particularly the consideration of effects and costs on multiple sectors beyond health where opportunity costs are still relevant. Despite advances in this area of research, there remain misconceptions about CE thresholds, the assumptions involved and their implications, which this paper aimed to highlight. It is the responsibility of all of us to educate those who are involved in priority setting about these concepts of threshold in order to ensure efficient healthcare resource allocation.

Acknowledgements The authors would like to thank Brian Reddy and Simon Walker for their advice on the Irish threshold and economic evaluation across multiple sectors, respectively.

Author contributions PT conceived the initial idea for the paper, and all authors (PT, JO, AL, TT) contributed to developing the outline of the paper. TT led on the hypothetical example, including the development of the Microsoft Excel file in the ESM; JO led on the sections on empirical estimates of the supply-side threshold; AL led on the section about the WHO-CHOICE guidelines and contributed to the demand side thresholds; and PT led on the remaining sections in the paper. All authors were involved in the multiple revisions of the paper before signing off on the final version.

Funding This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Compliance with ethical standards

Conflict of interest Praveen Thokala, Jessica Ochalek, Ashley Leech, and Thaison Tong have no conflicts of interest to declare.

References

1. Neumann PJ, Thorat T, Zhong Y, Anderson J, Farquhar M, Salem M, et al. A systematic review of cost-effectiveness studies reporting cost-per-DALY averted. PloS One. 2016;11(12):e0168512.
2. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. 4th ed. Oxford: Oxford University Press; 2015. p. 464.
3. Guide to the methods of technology appraisal 2013. Guidance and guidelines. NICE. https://www.nice.org.uk/process/pmg9/chapter/foreword. Accessed 14 Dec 2017.
4. Guidelines for the Economic Evaluation of Health Technologies: Canada. CADTH.ca https://www.cadth.ca/about-cadth/how-we-do-it/methods-and-guidelines/guidelines-for-the-economic-evaluation-of-health-technologies-canada. Accessed 14 Dec 2017.
5. Australian Government Department of Health. Guidelines for preparing a submission to the Pharmaceutical Benefits Advisory Committee: Version 5.0. 2016. https://pbac.pbs.gov.au/content/information/files/pbac-guidelines-version-5.pdf. Accessed Dec 2017.
6. Sculpher M, Claxton K, Pearson SD. Developing a value framework: the need to reflect the opportunity costs of funding decisions. Value Health. 2017;20(2):234–9.
7. Culyer T. Cost-effectiveness thresholds in health care: a bookshelf guide to their meaning and use. York: Centre for Health Economics, University of York; 2015. p. 1–22.
8. Bertram MY, Lauer JA, De Joncheere K, Edejer T, Hutubessy R, Kieny M-P, et al. Cost-effectiveness thresholds: pros and cons. Bull World Health Organ. 2016;94(12):925–30.
9. Woods B, Revill P, Sculpher M, Claxton K. Country-level cost-effectiveness thresholds: initial estimates and the need for further research. Value Health. 2016;19(8):929–35.

Journal : Large 40273 | Dispatch : 19-1-2018 | Pages : 14
Article No. : 606 | ☐ LE | ☐ TYPESET
MS Code : PECA-D-17-00309 | ☑ CP | ☑ DISK

10. Vallejo-Torres L, García-Lorenzo B, Castilla I, Valcárcel-Nazco C, García-Pérez L, Linertová R, et al. On the estimation of the cost-effectiveness threshold: Why, What, How? Value Health. 2016;19(5):558–66.
11. Schaffer SK, Cubi-Molla P, Devlin N, Towse A. Shaping the research agenda to estimate relevant cost-effectiveness thresholds for health technology assessment decision making: Report for ABPI. 2016. https://www.ohe.org/system/files/private/publications/Shaping%20the%20research%20agenda%20to%20estimate%20cost-effectiveness%20thresholds%20FOR%20PUBLICATION.pdf.
12. Crown W, Buyukkaramikli N, Thokala P, Morton A, Sir MY, Marshall DA, et al. Constrained optimization methods in health services research-an introduction: Report 1 of the ISPOR optimization methods emerging good practices task force. Value. 2017;20(3):310–9.
13. Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-effectiveness in health and medicine. Oxford, New York: Oxford University Press; 1996. p. 456.
14. Paulden M, O'Mahony J, McCabe C. Determinants of change in the cost-effectiveness threshold. Med Decis. 2017;37(2):264–76.
15. Stinnett AA, Paltiel AD. Mathematical programming for the efficient allocation of health care resources. J Health Econ. 1996;15(5):641–53.
16. Epstein DM, Chalabi Z, Claxton K, Sculpher M. Efficiency, equity, and budgetary policies: informing decisions using mathematical programming. Med Decis Mak. 2007;27(2):128–37.
17. Weinstein M, Zeckhauser R. Critical ratios and efficient allocation. J Public Econ. 1973;2(2):147–57.
18. Cleemput I, Neyt M, Thiry N, De Laet C, Leys M. Using threshold values for cost per quality-adjusted life-year gained in healthcare decisions. Int J Technol Assess Health Care. 2011;27(1):71–6.
19. Hummel JM, Oliveira MD, Bana e Costa CAB, IJzerman MJ. Supporting the project portfolio selection decision of research and development investments by means of multi-criteria resource allocation modelling. In: Marsh K, Goetghebeur M, Thokala P, Baltussen R (eds) Multi-criteria decision analysis to support healthcare decisions. Cham: Springer; 2017 [cited 14 Dec 2017]. pp. 89–103. https://link-springer-com.ezproxy.library.tufts.edu/chapter/10.1007/978-3-319-47540-0_6.
20. Eckermann S, Pekarsky B. Can the real opportunity cost stand up: displaced services, the straw man outside the room. PharmacoEconomics. 2014;32(4):319–25.
21. Hutubessy R, Chisholm D, Edejer TTT. Generalized cost-effectiveness analysis for national-level priority-setting in the health sector. Cost Eff Resour Alloc. 2003;1(1):8.
22. Newall AT, Jit M, Hutubessy R. Are current cost-effectiveness thresholds for low- and middle-income countries useful? Examples from the world of vaccines. PharmacoEconomics. 2014;32(6):525–31.
23. Phoya A, Araru T, Kachala R, Chizonga J, Bowie C. Disease control priorities in developing countries, 3rd edn. Working Paper #9: Setting Strategic Health Sector Priorities in Malawi. http://dcp-3.org/sites/default/files/resources/DCP%20Working%20Paper%209_Malawi%20Case%20Study_0.pdf.
24. McCabe C, Claxton K, Culyer AJ. The NICE cost-effectiveness threshold. PharmacoEconomics. 2008;26(9):733–44.
25. O'Mahony JF, Coughlan D. The Irish cost-effectiveness threshold: does it support rational rationing or might it lead to unintended harm of Ireland's health system? Value Health. 2015;18(7):A570.
26. Schwarzer R, Rochau U, Saverno K, Jahn B, Bornschein B, Muehlberger N, et al. Systematic overview of cost-effectiveness thresholds in ten countries across four continents. J Comp Eff Res. 2015;4(5):485–504.
27. Devlin N, Parkin D. Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. Health Econ. 2004;13(5):437–52.
28. Raftery JP. NICE's cost-effectiveness range: should it be lowered? PharmacoEconomics. 2014;32(7):613–5.
29. National Institute for Health and Care Excellence. Methods for the development of NICE public health guidance. (NICE process and methods guides). London: National Institute for Health and Care Excellence (NICE); 2012. http://www.ncbi.nlm.nih.gov/books/NBK395862/.
30. National Institute for Health and Clinical Excellence (NICE). National Institute for Health and Clinical Excellence: appraising life-extending, end of life treatments. 2009. https://www.nice.org.uk/guidance/gid-tag387/resources/appraising-life-extending-end-of-life-treatments-paper2.
31. Claxton K, Sculpher M, Palmer S, Culyer AJ. Causes for concern: is NICE failing to uphold its responsibilities to all NHS patients? Health Econ. 2015;24(1):1–7.
32. Framework Agreement between the Irish Pharmaceutical Healthcare Association Ltd and the Department of Health and the Health Service Executive on the Supply Terms, Conditions and Prices of Medicines [cited 14 Dec 2017]. http://www.hse.ie/eng/about/Who/cpu/IPHA_Agreement_2012.pdf.
33. NCPE Submission Process. National Centre for Pharmacoeconomics [cited 14 Dec 2017]. http://www.ncpe.ie/submission-process/.
34. Neumann PJ, Cohen JT, Weinstein MC. Updating cost-effectiveness: the curious resilience of the $50,000-per-QALY threshold. N Engl J Med. 2014;371(9):796–7.
35. Final Value Assessment Framework: updates for 2017–2019—ICER [cited 17 Oct 2017]. https://icer-review.org/material/final-vaf-2017-2019/.
36. Sullivan SD, Yeung K, Vogeler C, Ramsey SD, Wong E, Murphy CO, et al. Design, implementation, and first-year outcomes of a value-based drug formulary. J Manag Care Spec Pharm. 2015;21(4):269–75.
37. Leech AA, Kim D, Cohen J, Neumann PJ. Use and misuse of cost-effectiveness analysis thresholds in low and middle-income countries: trends in cost-per-DALY studies. Value Health. **(In press)**
38. WHO Commission on Macroeconomics and Health, World Health Organization. Macroeconomics and health: investing in health for economic development: executive summary. Macroéconomie et santé: investir dans la santé pour le développement économique: résumé d' orientation du rapport. 2001 [cited 17 Oct 2017]. http://www.who.int/iris/handle/10665/42463.
39. Robinson LA, Hammitt JK, Chang AY, Resch S. Understanding and improving the one and three times GDP per capita cost-effectiveness thresholds. Health Policy Plan. 2017;32(1):141–5.
40. Shillcutt SD, Walker DG, Goodman CA, Mills AJ. Cost-effectiveness in low- and middle-income countries. PharmacoEconomics. 2009;27(11):903–17.
41. Ochalek J. Cost per DALY averted thresholds for low- and middle-income countries—Research Database, The University of York. https://pure.york.ac.uk/portal/en/publications/cost-per-daly-averted-thresholds-for-low-and-middleincome-countries(12487fa5-e63f-4ac3-9fa4-03b2795065eb).html. Accessed 3 July 2017.
42. Claxton K, Martin S, Soares M, Rice N, Spackman E, Hinde S, et al. Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. Health Technol Assess. 2015;19(14):1–503 **(v–vi)**.
43. Marseille E, Larson B, Kazi DS, Kahn JG, Rosen S. Thresholds for the cost-effectiveness of interventions: alternative approaches. Bull World Health Organ. 2015;93(2):118–24.
44. Paris V, Belloni A. Value in pharmaceutical pricing. Country Profile: Australia. OECD. 2014 Nov (Working Paper No. 63).

Journal : Large 40273 | Dispatch : 19-1-2018 | Pages : 14
Article No. : 606 | ☐ LE | ☐ TYPESET
MS Code : PECA-D-17-00309 | ☑ CP | ☑ DISK

https://www.oecd.org/health/Value-in-Pharmaceutical-Pricing-Australia.pdf. Accessed Dec 2017.
45. Henry DA, Hill SR, Harris A. Drug prices and value for money: the Australian pharmaceutical benefits scheme. JAMA. 2005;294(20):2630–2.
46. Griffiths EA, Vadlamudi NK. Cadth's $50,000 cost-effectiveness threshold: fact or fiction? 2016; ISPOR. https://www.ispor.org/research_pdfs/54/pdffiles/PHP278.pdf. Accessed Dec 2017.
47. Grocott R, Metcalfe S, Alexander P, Werner R. Assessing the value for money of pharmaceuticals in New Zealand: PHARMAC's approach to cost-utility analysis. N Z Med J. 2013;126(1378):60–73.
48. O'Keeffe KM, Gander PH, Scott WG, Scott HM. Insomnia treatment in New Zealand. N Z Med J. 2012;125(1349):46–59.
49. Simoens S. Health economic assessment: a methodological primer. Int J Environ Res Public Health. 2009;6(12):2950–66.
50. Metcalfe S, Rodgers A, Werner R, Schousboe C. PHARMAC has no cost-effectiveness threshold. N Z Med J. 2012;125(1350):99–101.
51. Metcalfe S, Grocott R. Comments on "Simoens, S. Health economic assessment: a methodological primer. Int. J. Environ. Res. Public Health 2009, 6, 2950–2966"—New Zealand in fact has no cost-effectiveness threshold. Int J Environ Res Public Health. 2010;7(4):1831–4.
52. Scottish Medicines Consortium. SMC Modifiers used in Appraising New Medicines. https://www.scottishmedicines.org.uk/About_SMC/Policy_statements/SMC_Modifiers_used_in_Appraising_New_Medicines. Accessed 14 Dec 2017.
53. Kamae I. Value-based approaches to healthcare systems and pharmacoeconomics requirements in Asia: South Korea, Taiwan, Thailand and Japan. Pharmacoeconomics. 2010;28(10):831–8.
54. Coelho De Soarez P, Dutilh Novaes HM. Cost-effectiveness thresholds and the Brazilian Unified National Health System. 2017. http://www.scielo.br/pdf/csp/v33n4/1678-4464-csp-33-04-e00040717.pdf. Accessed Dec 2017.
55. Gravelle HS, Backhouse ME. International cross-section analysis of the determination of mortality. Soc Sci Med. 1987;25(5):427–41.
56. Barnsley P, Towse A, Karlsberg Schaffer S, Sussex J. Critique of CHE Research Paper 81: methods for the estimation of the NICE cost effectiveness threshold. OHE. 2013. https://www.ohe.org/publications/critique-che-research-paper-81-methods-estimation-nice-cost-effectiveness-threshold. Accessed 14 Dec 2017.
57. Claxton K, Sculpher M. Response to the OHE critique of CHE Research paper 81. https://www.york.ac.uk/media/che/documents/Response%20to%20the%20OHE%20critique%20of%20CHE%20Research%20paper%2081.pdf. Accessed Dec 2017.
58. Edney L, Afzali H, Cheng T, Karnon J. Estimating the reference ICER for the Australian health system. Pharmacoeconomics. 2017. https://doi.org/10.1007/s40273-017-0585-2 (**Epub 22 Dec 2017**).
59. Vallejo-Torres L, GarcÌa-Lorenzo B, Serrano-Aguilar P. Estimating a cost-effectiveness threshold for the Spanish NHS. FEDEA; 2016 Jun. Report No. eee2016-22. https://www.ucl.ac.uk/dahr/pdf/HESG/Paper_A31.pdf. Accessed Dec 2017.
60. Government of the Republic of Malawi. Health Sector Strategic Plan II (2017–2022): towards Universal Coverage. 2017. file:///Users/ashleyleech/Downloads/HSSP%20II%20Final%20HQ%20complete%20file.pdf.pdf. Accessed Dec 2017.
61. Ochalek J, Claxton K, Revill P, Sculpher M, Rollinger A. Supporting the development of an essential health package: principles and initial assessment for Malawi. Centre for Health Economics, University of York; 2016 Sep. (Working Papers). Report No.: 136cherp https://www.york.ac.uk/media/che/documents/papers/researchpapers/CHERP136_EHP_Malawi_interventions.pdf. Accessed 17 Oct 2017.
62. Thavorncharoensap M, Teerawattananon Y, Natanant S, Kulpeng W, Yothasamut J, Werayingyong P. Estimating the willingness to pay for a quality-adjusted life year in Thailand: does the context of health gain matter? Clinicoecon Outcomes Res. 2013;5:29–36.
63. Teerawattananon Y, Tritasavit N, Suchonwanich N, Kingkaew P. The use of economic evaluation for guiding the pharmaceutical reimbursement list in Thailand. Z Evidenz Fortbild Qual Im Gesundheitswesen. 2014;108(7):397–404.
64. Lim YW, Shafie AA, Chua GN, Ahmad Hassali MA. Determination of cost-effectiveness threshold for health care interventions in Malaysia. Value Health. 2017;20(8):1131–8.
65. GDP per capita (current US$). Data. https://data.worldbank.org/indicator/NY.GDP.PCAP.CD. Accessed 14 Dec 2017.
66. Initiative on US Value Assessment Frameworks. https://www.ispor.org/ValueAssessmentFrameworks/Index. Accessed 17 Oct 2017.
67. Sanders GD, Neumann PJ, Basu A, Brock DW, Feeny D, Krahn M, et al. Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: second panel on cost-effectiveness in health and medicine. JAMA. 2016;316(10):1093–103.
68. Neumann PJ, Sanders GD, Russell LB, Siegel JE, Ganiats TG, editors. Cost-effectiveness in health and medicine. 2nd ed. New York: Oxford University Press; 2016. p. 536.
69. Devlin NJ, Sussex J. Incorporating multiple criteria in HTA: methods and processes. OHE, Office of Health Economics Research; 2011. https://healthpolicy.fsi.stanford.edu/sites/default/files/ohe_hta_methods.pdf. Accessed Dec 2017.
70. Carrera P, Ijzerman MJ. Are current ICER thresholds outdated? Valuing medicines in the era of personalized healthcare. Expert Rev Pharmacoecon Outcomes Res. 2016;16(4):435–7.
71. Marsh K, Ijzerman M, Thokala P, Baltussen R, Boysen M, Kaló Z, et al. Multiple criteria decision analysis for health care decision making: emerging good practices: report 2 of the ISPOR MCDA Emerging Good Practices Task Force. Value Health. 2016;19(2):125–37.
72. Remme M, Martinez-Alvarez M, Vassall A. Cost-effectiveness thresholds in global health: taking a multisectoral perspective. Value Health. 2017;20(4):699–704.
73. Nimdet K, Chaiyakunapruk N, Vichansavakul K, Ngorsuraches S. A systematic review of studies eliciting willingness-to-pay per quality-adjusted life year: does it justify CE threshold? PloS One. 2015;10(4):e0122760.
74. Gafni A, Birch S. Incremental cost-effectiveness ratios (ICERs): the silence of the lambda. Soc Sci Med. 2006;62(9):2091–100.
75. Afzali HHA, Karnon J, Sculpher M. Should the Lambda (λ) Remain Silent? Pharmacoeconomics. 2016;34(4):323–9.
76. Towse A. Should NICE's threshold range for cost per QALY be raised? Yes. BMJ. 2009;338:b181.
77. Collins M, Latimer N. NICE's end of life decision making scheme: impact on population health. BMJ. 2013;346:f1363.
78. Lomas J, Claxton K, Martin S, Soares M. Resolving the "cost-effective but unaffordable" "paradox": estimating the health opportunity costs of non-marginal budget impacts. Value Health. (**In Press**).
79. Chhatwal J, Kanwal F, Roberts MS, Dunn MA. Cost-effectiveness and budget impact of hepatitis C virus treatment with sofosbuvir and ledipasvir in the United States. Ann Intern Med. 2015;162(6):397–406.
80. National Institute for Health and Care Excellence. Technology Appraisal and Highly Specialised Technologies Programmes. Procedure for varying the funding requirement to take account of net budget impact. https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/

Journal : Large 40273 | Dispatch : 19-1-2018 | Pages : 14
Article No. : 606 | LE | TYPESET
MS Code : PECA-D-17-00309 | CP | DISK

TA-HST-procedure-varying-the-funding-direction.pdf. Accessed 14 Dec 2017.

81. Ubel PA, Hirth RA, Chernew ME, Fendrick AM. What is the price of life and why doesn't it increase at the rate of inflation? Arch Intern Med. 2003;163(14):1637–41.
82. Thokala P, Dixon S, Jahn B. Resource modelling: the missing piece of the HTA jigsaw? Pharmacoeconomics. 2015;33(3):193–203.
83. Vassall A, Mangham-Jefferies L, Gomez GB, Pitt C, Foster N. Incorporating demand and supply constraints into economic evaluations in low-income and middle-income countries. Health Econ. 2016;25:95–115.
84. Griffiths EA, Hendrich JK, Stoddart SD, Walsh SC. Acceptance of health technology assessment submissions with incremental cost-effectiveness ratios above the cost-effectiveness threshold. Clinicoecon Outcomes Res. 2015;7:463–76.

Journal : Large 40273 | Dispatch : 19-1-2018 | Pages : 14
Article No. : 606 | □ LE | □ TYPESET
MS Code : PECA-D-17-00309 | ✔ CP | ✔ DISK