

1 LEADING ARTICLE

2 **Cost-Effectiveness Thresholds: the Past, the Present**
3 **and the Future**


4 Praveen Thokala¹  · Jessica Ochalek² · Ashley A. Leech³ · Thaison Tong¹

5
6 © Springer International Publishing AG, part of Springer Nature 2018

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

26
27

A1 **Electronic supplementary material** The online version of this
A2 article (<https://doi.org/10.1007/s40273-017-0606-1>) contains supple-
A3 mentary material, which is available to authorized users.

A4  Praveen Thokala
A5 p.thokala@sheffield.ac.uk

A6 ¹ Health Economics and Decision Science, School of Health
A7 and Related Research, University of Sheffield, Regent Court,
A8 30 Regent Street, Sheffield S1 4DA, UK

A9 ² Centre for Health Economics, University of York,
A10 Heslington, York YO10 5DD, UK

A11 ³ Center for the Evaluation of Value and Risk in Health
A12 (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

31
32
29
33
34
35
36
37
38
39
40
41
42
43
44
46
47

48

49
50
51
52
53
54
55
56
57
58
59

60 identify whether the new intervention is good VfM. Inter-
61 ventions with an ICER below a threshold (i.e. if they add
62 each QALY or avert each DALY at a lower cost than the
63 threshold) are considered cost effective, while those with
64 an ICER above the threshold are not.¹

65 Despite the widespread recognition and use of CE
66 thresholds by researchers conducting economic evaluations
67 [1, 2], as well as the adoption of their use into policy in
68 some countries (e.g. the UK, Australia and Canada) [3–6],
69 there is a lack of understanding among many about the
70 meaning of thresholds, the assumptions involved, and their
71 implications. Whether a given intervention is cost-effective
72 or not depends on how much health it would generate and
73 whether that amount is greater than the health that could
74 have been generated if the money required to fund it had
75 been spent on something else, which is a measure of
76 opportunity cost. As such, using a CE threshold to reflect
77 this perspective has come to be known as a ‘supply side’
78 approach [7–10]. When non-health impacts on private
79 consumption are also considered important, some assess-
80 ment of the equivalent consumption value of health is
81 required, i.e. ‘demand side’ empirical research [10]. Such
82 approaches aim to represent societal willingness to pay
83 (WTP) for additional health gains, i.e. what individuals are
84 willing to forego in non-healthcare/private consumption for
85 gains in healthcare. When considering budget constraints
86 on a healthcare system, supply-side thresholds can be
87 considered more relevant since displacements to current
88 health-generating interventions must happen to fund new
89 interventions [11].

90 The aim of this paper is threefold: (1) to provide an
91 illustration of the CE threshold using a hypothetical
92 example to highlight the key assumptions involved; (2) to
93 describe the various thresholds that are in use as policy
94 tools in countries or have been estimated by researchers
95 (sometimes, though not always overlapping—see the
96 example of the UK); and (3) to present the new develop-
97 ments and ongoing areas of research around thresholds.
98 The remainder of the paper is structured as follows. Sec-
99 tion 2 presents a simple hypothetical example to illustrate
100 how the CE threshold can be determined using the ‘league
101 table’ approach, as well as optimization techniques and the
102 assumptions involved. Section 3 describes the use of his-
103 toric/heuristic estimates of the threshold, along with their
104 limitations. Then, in Sect. 4, recent endeavours to estimate
105 the empirical value of thresholds (e.g. work on opportunity
106 costs in UK, Australia, Spain, as well as work in estimating
107 thresholds for low- to middle-income countries [LMICs])
108 will be presented. Section 5 presents future directions for

the field (thresholds across sectors such as social care, 109
incorporating multiple criteria beyond QALYs) and their 110
impact on CE thresholds. Finally, Sect. 6 addresses some of 111
the common issues and misconceptions associated with CE 112
thresholds. 113

2 An (Hypothetical) Example 114

In this section, a simple hypothetical example is used to 115
illustrate how the CE threshold can be determined using a 116
‘league table’ approach and optimization techniques [12]. 117
We further outline the assumptions involved in each. Let us 118
assume there is a fixed healthcare budget of £50 million 119
available and the aim is to choose interventions to place in 120
the healthcare package to maximize the total QALYs 121
gained. In this example, for the sake of simplicity, the 122
healthcare package is empty to start with and there are 123
seven mutually exclusive, independent interventions to 124
choose from, each with a different set of costs and QALYs 125
gained, as shown in Table 1. Note that these are incre- 126
mental costs and QALYs associated with each intervention 127
compared with the ‘do nothing’ option. At first glance, it is 128
obvious that the budget of £50 million is not enough to 129
fund all interventions. 130

2.1 League Table Approach 131

Under certain assumptions, which are outlined below, this 132
‘league table’ approach [13] can be used to identify the 133
optimal allocation by including interventions according to 134
highest VfM until the available budget is exhausted. Given 135
our aim is to maximize health, the measure of ‘value’ in 136
our example is QALYs. As we started with an empty 137
package and are considering only independent options, we 138
calculate VfM by dividing the costs by the QALYs of each 139
intervention, as presented in the fourth column of Table 2 140
(i.e. they represent the ICERs for each intervention 141

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

1FL01 ¹ CE thresholds reflecting opportunity costs can also be used to
1FL02 calculate the net benefit of an intervention (i.e. if net health benefit,
1FL03 benefit in terms of health over and above health opportunity costs).

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

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

148 As shown in Table 2, intervention A has a ranking of 1
 149 (i.e. provides the best VfM), therefore it is added to the
 150 healthcare package first. Intervention A generates 7000
 151 QALYs at the rate of £4571/QALY, with a total cost of
 152 £32 million, therefore there is £18 million still left from
 153 the overall budget of £50 million. The next best interven-
 154 tion is D, which costs £10 million and provides 2000
 155 QALYs at the rate of £5000/QALY. After incorporating
 156 intervention D into the healthcare package, there is
 157 £8 million still left, which can be spent on the next best
 158 intervention, B. However, £8 million is not enough to fund
 159 intervention B in full (with a cost of £22 million), there-
 160 fore, we can only fund a portion (8 million/£22 mil-
 161 lion = 0.36) within the budget. This would result in a gain
 162 of 1454.5 QALYs (i.e. 0.36×4000 QALYs) from inter-
 163 vention B at the rate of £5500/QALY. In total, we achieved
 164 10,454.5 QALYs (7000 QALYs from A, 2000 QALYs
 165 from D, and 1454.5 QALYs from B) for the £50 million
 166 budget (see the Microsoft Excel file in the electronic sup-
 167plementary material [ESM] for a visual illustration of this
 168 approach as a ‘bookshelf’) [7, 14].

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

175 **2.2 Budget-Constrained Optimization**

176 Mathematical programming techniques can also be used to
 177 identify the optimal allocation that maximizes the total
 178 QALYs gained within the budget constraint [15, 16] (see

ESM for the solution of the budget-constrained optimiza- 179
 tion problem). It can be seen that the optimal solution 180
 achieved is the same as that found using the league 181
 table approach. However, these two approaches find the 182
 same result only under a strict set of assumptions (perfect 183
 divisibility, linearity, and independence), which are 184
 described later in the ‘Underlying Assumptions’ section. 185

186 If the budget is bigger, say £51 million, we could gain a 186
 further 181.8 QALYs by spending the additional £1 mil- 187
 lion on intervention B. In fact, at the current allocation of 188
 the £50 million budget (A, B, and D), 0.0001818 additional 189
 QALYs can be gained for every £1 increase in the budget. 190
 In optimization terminology, this is termed the shadow 191
 price, i.e. how much the objective (QALYs) would increase 192
 for a one-unit increase in the constraint (budget). The 193
 shadow price can also be presented as decrements, i.e. how 194
 much the objective (QALYs) would decrease for a one-unit 195
 decrease in the constraint (budget). In our example, 196
 0.0001818 is the shadow price of the £50 million budget 197
 optimally allocated. It should be noted that this shadow 198
 price is the inverse of the cost per QALY of the last 199
 intervention included (£5500 per QALY for intervention 200
 B). Also note that this shadow price is only applicable for a 201
 range of budget between £42 million (i.e. total costs of 202
 fully funded A and D) and £64 million (i.e. total costs of 203
 fully funded A, D, and B). 204

205 The inverse of the shadow price at the optimal allocation 205
 in the budget, referred to as the ‘critical ratio’ in one of 206
 the first mentions of the threshold in published literature [17], 207
 represents the ‘supply side’ definition of the CE threshold, 208
 i.e. a threshold representing the notion of opportunity cost. 209
 Whether a given intervention is cost effective or not thus 210
 depends on how much health it would generate and whe- 211
 ther that amount is greater than the health that could have 212
 been generated if the money required to fund it had been 213
 spent on something else, which is a measure of opportunity 214
 cost. 215

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

216 2.3 Assessing Cost-Effectiveness

217 With this allocation of the £50 million budget, for a new
 218 intervention X to be included in the healthcare package (X
 219 does not have to be from the existing list in Table 1), we
 220 would need to disinvest first (assuming the overall budget
 221 is fixed at £50 million). This disinvestment is only worth-
 222 while if the replacement of existing interventions in the
 223 current healthcare package with X brings positive net
 224 QALYs gained. Let us introduce a new intervention X that
 225 costs £5.2 million and provides 1000 QALYs at the rate of
 226 £5200 per QALY (Table 3). Given the existing allocation
 227 (A, B and D), the decision is whether we should fund X.

228 As illustrated earlier, the threshold is the inverse of the
 229 shadow price of the budget with its current optimal allo-
 230 cation, which is £5500 per QALY (i.e. the cost per QALY
 231 of the last intervention included, intervention B). Since
 232 £5200/QALY (VfM of X) is lower than £5500/QALY
 233 (current inverse of the shadow price), it is cost effective to
 234 replace B with X. That is, more QALYs can be gained by
 235 spending money on X than those lost by displacing part of
 236 B. In this case, X will be funded from the replacement of
 237 part of B, the proportion of B left after funding X is esti-
 238 mated as follows: $(£8 - 5.2 \text{ million}) / £22 \text{ million} = 12.7\%$.
 239 Replacing B with X would generate 1509.1 QALYs for the
 240 £8 million (i.e. 1000 QALYs from X + $0.127 * 4000$ QALYs
 241 from part of B). In total, we achieve 10,509.1 QALYs (7000
 242 QALYs from A, 2000 QALYs from D, 1000 QALYs from
 243 X, and 509.1 QALYs from B) for the £50 million budget, an
 244 increase of 54.6 QALYs ($10,509.1 - 10,454.5 = 54.6$
 245 QALYs) compared with the previous allocation.

246 2.4 Underlying Assumptions

247 Through this example, we illustrate below a few key
 248 assumptions relating to CE thresholds that are worth fur-
 249 ther consideration, i.e. perfect divisibility, linearity, inde-
 250 pendence, marginal increments in budget, disinvestment
 251 plan, perfect information and other issues [14, 18].

252 2.4.1 Perfect Divisibility, Linearity and Independence

253 One assumption that applies to both the league table ap-
 254 proach 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 intro-
 duced), 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
 $(0.36 = £8 \text{ million} / £22 \text{ million})$ 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 (be-
 fore 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 opti-
 mization problem can be structured using integer pro-
 gramming to overcome these issues, the shadow prices are

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

300 *2.4.2 Marginal Budget Impact*

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

326 *2.4.3 Disinvestment Plan*

327 In our example, we assume that the disinvestment to fund
328 a new intervention should come from the least cost-ef-
329 fective intervention(s). The new intervention was only
330 compared with the least cost effective existing interven-
331 tion within the optimal allocation, to keep with our
332 original aim of maximizing QALYs. Replacing interven-
333 tions other than the least cost-effective intervention (i.e.
334 anything other than the intervention with least VfM) in
335 our healthcare package will result in greater QALYs lost
336 than when displacing least cost-effective intervention.
337 However, it is not always possible to ensure that the least
338 cost-effective intervention(s) are disinvested first or that
339 the healthcare package is ‘optimal’ [20]. Healthcare
340 packages in real-life settings tend to include a mix of
341 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 relation-
ship 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 impli-
cations 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 opti-
mization 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 sec-
tion, many countries use a threshold value based on other
methods and representing different concepts. For exam-
ple, 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/judge-
ment [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 esti-
mates, whether explicit or implied, are used across the
world.

2FL01 ² As described earlier, the shadow price of 0.0001818 relates to
2FL02 intervention B and as such is only applicable for a range of budgets
2FL03 between £42 million (i.e. total costs of fully funded A and D) and
2FL04 £64 million (i.e. total costs of fully funded A, D and B).

388	3.1 Explicit Thresholds		
389	<i>3.1.1 UK (National Institute for Health and Care</i>		
390	<i>Excellence)</i>		
391	The National Institute for Health and Care Excellence		
392	(NICE) in the UK is a high-profile example of the use of		
393	explicit CE thresholds, and its guidance recommends in		
394	favour of funding interventions with an ICER below a		
395	threshold of £20,000/QALY or £30,000/QALY, and also		
396	recommends against funding interventions with an ICER		
397	above these thresholds [27–29]. However, a higher		
398	threshold (i.e. £50,000/QALY) is used for life-extending		
399	treatments for small patient populations at the end of life,		
400	i.e. treatments that offer an extension to life greater than		
401	3 months compared with current treatment in the National		
402	Health Service (NHS); are for patients with a short life		
403	expectancy, i.e. normally <24 months; and are for small		
404	patient populations, normally not exceeding a cumulative		
405	total of 7000 patients for all licensed indications in England		
406	[30]. Despite this guidance, interventions with ICERs		
407	above £30,000 or £50,000 are often accepted, even when		
408	lacking the requisite special evidence needed [31].		
409	<i>3.1.2 Ireland</i>		
410	The CE of all new medicines in Ireland is considered by the		
411	National Centre for Pharmacoeconomics (NCPE), in col-		
412	laboration with the Health Service Executive (HSE), the		
413	public body with responsibility for delivering state-funded		
414	healthcare in Ireland. The Irish Pharmaceutical Healthcare		
415	Association (IPHA) and HSE have an agreement that		
416	explicitly states that the QALY threshold to be used in the		
417	HTA process is €45,000 [32]. This value is also confirmed		
418	on the NCPE website [33]. It is worth noting that, unlike		
419	NICE, NCPE's recommendations are not mandatory and		
420	can be overruled by the minister/HSE [25].		
421	<i>3.1.3 US</i>		
422	While \$50,000 per QALY has been mentioned anecdotally		
423	in the past in the US [34], the recent value frameworks		
424	mention explicit thresholds. Given the diversity of payers		
425	and healthcare organizations, it should be noted that there		
426	are differences in the thresholds used. A high-profile		
427	example of explicit reference to thresholds is the use of		
428	\$100,000–\$150,000/QALY for a value-based price		
429	benchmark by the Institute for Clinical and Economic		
430	Review (ICER) [35], a trusted non-profit organization that		
431	evaluates evidence on new technologies in the US. Premera		
432	Blue Cross, a large not-for-profit health plan in the Pacific		
433	Northwest, uses value-based formulary tiers based on		
434	ICER thresholds—drugs are allocated to one of the four co-		
	payment tiers (tier 1, <\$10,000/QALY; tier 2, \$10,000 to	435	
	<\$50,000/QALY; tier 3, \$50,000 to <\$150,000/QALY;	436	
	and tier 4, >\$150,000/QALY) [36].	437	
	3.2 Heuristics for the Threshold Value: WHO-	438	
	CHOICE (One to Three Times a Country's	439	
	Gross Domestic Product)	440	
	One to three times a country's annual GDP per capita has	441	
	been a widely used threshold for CE studies within global	442	
	health, mainly among studies focused on LMICs [1, 37]. A	443	
	recent study found that the proportion of LMICs citing this	444	
	threshold has substantially increased over time, with 10%	445	
	of studies citing this threshold in the early 2000s, to 76%	446	
	between 2013 and 2015 [37]. While the origins for its	447	
	intended use for CEA are less clear, the WHO first used	448	
	these values in its 2001 Commission on Macroeconomics	449	
	and Health (CMH) report [38]. While this report intuitively	450	
	equates a year of life to per capita income, considering	451	
	productivity and leisure time, it used per capita income to	452	
	value the economic loss resulting from the burden of major	453	
	diseases impacting countries. Despite its variant aim, the	454	
	WHO-CHOICE thereafter adopted this range for promot-	455	
	ing CEA [21, 38]. There have recently been several opin-	456	
	ions on this threshold value that have motivated calls for	457	
	consensus and new primary research [8, 9, 37, 39–43]. For	458	
	instance, some analysts have argued that CE thresholds	459	
	reflecting opportunity costs are much lower than the one to	460	
	three times GDP per capita rule of thumb, while other	461	
	analysts encourage applying a range of income elasticity	462	
	estimates to account for the relationship between the value	463	
	per statistical life (VSL) and income [39]. The WHO has	464	
	since backed away from this threshold range and recog-	465	
	nizes its limitations for CEA [8].	466	
	3.3 Implied/Unspecified Thresholds	467	
	A recent systematic overview of CE thresholds suggested	468	
	that many countries do not specify a threshold [26]. While	469	
	researchers analysed previous decisions to identify the	470	
	threshold value in these countries, they were unable to pin	471	
	down a single number. Nevertheless, the manner in which	472	
	these countries use different CE thresholds is briefly	473	
	described below.	474	
	<i>3.3.1 Pharmaceutical Benefits Advisory Committee</i>	475	
	The Pharmaceutical Benefits Advisory Committee (PBAC)	476	
	in Australia does not formally specify a CE threshold.	477	
	However, the cost per QALY of the technology is reported	478	
	as belonging to one of four bands, i.e. AUS\$15,000–	479	
	\$45,000; \$45,000–\$75,000; \$75,000–\$105,000; \$105,000–	480	
	\$200,000. A recent study by Paris and Belloni [44] at the	481	

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

490 3.3.2 Canadian Agency for Drugs and Technologies

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

504 3.3.3 New Zealand

505 The Pharmaceutical Management Agency (PHARMAC) in
 506 New Zealand state that they do not have a CE threshold
 507 [47]. While researchers have tried to imply the threshold
 508 from previous decisions [48, 49], PHARMAC states that
 509 they fund medicines within a fixed budget, and as CE is
 510 only one of its nine decision criteria used to inform deci-
 511 sions, thresholds cannot be inferred or calculated [50].
 512 They also note that CE estimates for PHARMAC’s
 513 investments has ranged between – NZ\$40,000 (net cost
 514 savings to the health sector for health gains) to over
 515 + NZ\$200,000 per QALY (– €20,000 to + €100,000) [51].

516 3.3.4 Other Countries

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

520 **4 Present: Empirical Estimates of Cost-
 521 Effectiveness (CE) Threshold**

522 Recently, some countries have begun to conduct empirical
 523 research to identify CE thresholds for their setting. These
 524 studies have broadly been classified as either supply- or
 525 demand-side estimates [10]. Supply-side estimates aim to

reflect the opportunity cost of spending on health by link- 526
 ing the healthcare expenditure to health outcomes, while 527
 the demand-side estimates aim to reflect societal WTP for 528
 improvements in health. 529

4.1 Supply-Side Thresholds 530

It should be noted that the example in Sect. 2 illustrates an 531
 ideal situation in which the budget allocation is optimal; it 532
 is easy to identify the least cost-effective intervention(s), 533
 and the system (decision makers) only displace these least 534
 cost-effective interventions. This is a ‘first best’ situation; 535
 however, in practice, this is not always the case. In com- 536
 plex systems, the existing healthcare package may not be 537
 optimal, it may not be possible to specify exactly what 538
 activities are displaced, and decisions about disinvestment 539
 may be left to other decision makers in the system, for 540
 example at a local level. Thus, in empirically estimating 541
 the threshold, the aim is to estimate the shadow price of the 542
 budget in terms of the interventions that are likely to be 543
 displaced [42]. This is what Culyer describes as an 544
 approach to estimating the ‘second best’ threshold [7]. 545
 These empirical estimates of the supply-side threshold tend 546
 to reflect the marginal productivity of the healthcare sys- 547
 tem, derived from the relationship between changes in 548
 healthcare expenditure and health outcomes, where 549
 expenditures at the margin may be committed to a mix of 550
 cost-effective and cost-ineffective interventions (i.e. inter- 551
 ventions with a range of cost per QALYs) [7]. In a world 552
 where the assumptions of the optimization model are met, 553
 this conceptualization of the threshold should result in the 554
 same value as that which is derived by solving the con- 555
 strained optimization problem. However, where the nec- 556
 essary assumptions as set out in the preceding section are 557
 not met, the values may differ. The ‘second best’ approach 558
 provides an estimate that best informs the expected health 559
 opportunity costs of a new intervention and, therefore, if 560
 robustly estimated, can be better relied on to inform whe- 561
 ther a new intervention is expected to result in a net health 562
 gain or net health loss. 563

There are challenges involved in estimating the rela- 564
 tionship between changes in healthcare expenditure and 565
 health outcomes, i.e. the marginal productivity of the 566
 healthcare system. Given the outcome of interest is QALYs 567
 [a combination of quality of life (QoL) and life-years 568
 (LYs)], there is a need to link the healthcare expenditure to 569
 mortality (to estimate the effect on LYs) and morbidity (to 570
 estimate the effect on QoL). The data on healthcare 571
 expenditure and its effect on mortality/morbidity may not 572
 always be readily available and, as such, assumptions are 573
 often required. Furthermore, there are also econometric 574
 challenges that include, but are not limited to, issues 575
 around controlling for the many non-healthcare factors that 576

577	affect health [55], which if not properly accounted for may	622
578	lead to biased and inconsistent estimates. To date, such	623
579	within-country estimation has been undertaken in relatively	624
580	few countries, which are described below alongside one	625
581	example where cross-country data have been used to esti-	626
582	mate these values for a number of countries.	627
583	4.1.1 UK	628
584	Claxton et al. [42] empirically estimated the CE threshold	629
585	for the NHS in the UK to be £12,936 per QALY. They used	630
586	the English NHS programme budgeting data to estimate the	631
587	relationship between changes in overall NHS expenditure	632
588	and changes in mortality/LYs gained, and subsequently	633
589	extended this to QALYs. Their ‘structural’ uncertainty	634
590	suggested that the estimate is likely to be an overestimate	635
591	and reported that the probability the threshold is less than	636
592	£20,000 per QALY is 0.89, and the probability that it is less	637
593	than £30,000 per QALY is 0.97. The assumptions made in	
594	the estimation of the UK threshold have been discussed in a	
595	number of publications [28, 56, 57].	
596	4.1.2 Australia	638
597	Edney et al. [58] estimated the CE threshold, called the	639
598	reference ICER, for Australia. They used an instrumental	640
599	variable two-stage least squares regression to estimate the	641
600	effect of changes in health expenditure on QALYs due to	642
601	reduced mortality. Further empirical analysis was then used	643
602	to inform the effect of health expenditure in terms of	644
603	QALYs due to reduced morbidity. These are then com-	
604	bined to produce a central estimate of the reference ICER,	
605	which represents the average opportunity costs of decisions	
606	to fund new technologies, i.e. AUS\$28,033/QALY.	
607	4.1.3 Spain	
608	Vallejo-Torres et al. [59] estimated the CE threshold for	639
609	the Spanish NHS. They used 5 years of data across the 17	640
610	regional health services in Spain to regress quality-adjusted	641
611	life expectancy (QALE) against health spending, control-	642
612	ling for region and year fixed effects, and a comprehensive	643
613	set of time- and region-variant indicators, applying a 1-year	644
614	lag to expenditure. They report that health expenditure has	645
615	a positive and significant effect on QALE, with an average	646
616	spending elasticity of 0.07, which translates into a cost per	647
617	QALY of between €21,000 and €24,000.	648
618	4.1.4 Low- to Middle-Income Countries CE Thresholds	649
619	Ochalek [41] estimate CE thresholds for 123 LMICs using	650
620	estimates of the effect of a change in government spending	651
621	on health on health outcomes from cross-country data.	652
	Their study expands on existing studies within the litera-	653
	ture estimating the effect of a change in spending on	654
	mortality outcomes to estimate the effect of a change in	655
	spending on a range of mortality and morbidity outcomes.	656
	Using data on each country’s demography (i.e. the sex and	657
	age structure of the population), epidemiology (i.e.	658
	underlying mortality and morbidity burden) and health	659
	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 oppor-	
	tunity 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	660
	The empirical methods of estimating demand-side thresh-	661
	olds, namely WTP and value of a statistical life studies, are	662
	reviewed and discussed in detail by Vallejo-Torres et al.	663
	[10]. Below, we offer a brief description of the application	664
	of these methods in policy in two countries—Thailand and	665
	Malaysia.	666
	4.2.1 Thailand (Health Intervention and Technology	667
	Assessment Program)	668
	The Health Intervention and Technology Assessment Pro-	669
	gram (HITAP) in Thailand elicited the WTP for a QALY in	670
	the Thai healthcare setting [62]. The results of this study	671
	were adopted by decision-making bodies as the appropriate	672
	threshold for health investment in the Thai setting; the	673
	ceiling threshold is reported to be 160,000 Baht per QALY,	674
	which is approximately 1.2 times Gross National Income	675
	(GNI) per capita [63]. However, they also note that this	676
	single threshold is not used for resource allocation of all	677
	types of interventions; for example, sometimes medicines	678
	that treat rare diseases are included in the National List of	679
	Essential Medicines (NLEM) even though their ICER is	680
	much higher than the threshold.	681
	4.2.2 Malaysia	682
	Lim et al. conducted a cross-sectional, contingent valuation	683
	study in four states of Malaysia to estimate the CE	684
	threshold for healthcare interventions as WTP for a QALY	685
	[64]. One thousand and thirteen respondents were inter-	686
	viewed in person for their socioeconomic background,	687
	QoL, and WTP for a hypothetical scenario. The authors	688
	reported that the CE thresholds ranged from MYR12,810 to	689
	MYR28,470 (US\$4000–US\$8900) and education level,	690

669 estimated monthly household income, and the description
 670 of health state scenarios had the biggest effect on the WTP
 671 estimates. They concluded that there is no single WTP
 672 value for a QALY and that the CE threshold estimated for
 673 Malaysia was found to be lower than the threshold value
 674 recommended by the WHO (i.e. one and three times the
 675 GDP per capita, which was approximately \$10,000 and
 676 \$30,000, respectively, in 2017) [65].

677 **5 Future: Beyond Quality-Adjusted Life-Years**
 678 **(QALYs)? Other Sectors?**

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

690 **5.1 Thresholds for Benefits Beyond QALYs**

691 Alongside the recent work on value frameworks [66],
 692 which mentions many additional criteria beyond QALYs, it
 693 is widely acknowledged that many HTA organizations
 694 consider multiple factors alongside CE [69]. More recently,
 695 there have been calls for including these multiple criteria
 696 explicitly in the assessment of value [70], using techniques
 697 such as multicriteria decision analysis (MCDA) [71]. The
 698 current CE thresholds are based on QALYs (or DALYs)
 699 being the measure of effectiveness. If the value is redefined
 700 to include multiple criteria beyond QALYs (or DALYs),
 701 the measure of effectiveness is not QALYs (or DALYs)
 702 anymore but rather a new composite measure of effec-
 703 tiveness. As such, the threshold will need to be re-esti-
 704 mated for this new measure of ‘effectiveness’ to reflect the
 705 opportunity costs [6]. As observed in Sect. 4 (the empirical
 706 estimates of the supply-side thresholds), this poses a sig-
 707 nificant informational challenge in identifying the marginal
 708 impacts on the different criteria that make up the overall
 709 effectiveness.

710 **5.2 Thresholds in Other Sectors**

711 The Second Panel on Cost-Effectiveness in Health and
 712 Medicine [67, 68] supports a societal perspective and
 713 recommends the use of an ‘impact inventory’—a structured
 714 table listing the health and non-health effects of an

intervention that should be considered in a societal refer- 715
 ence-case analysis. To evaluate interventions crossing 716
 multiple sectors, sector-specific thresholds are needed that 717
 represent the sector-specific outcome that would be for- 718
 gone as the result of the additional costs of a new inter- 719
 vention. To date, no sector outside of healthcare has 720
 established a threshold. While some sectors have estab- 721
 lished measures, such as the Adult Social Care Outcomes 722
 Toolkit (ASCOT) used to estimate social care-related QoL 723
 (SCRQoL), many sectors do not have standard definitions 724
 for their outcomes. The challenges involved in performing 725
 CEA when the intervention concerns multiple sectors are 726
 highlighted by Remme et al. [72]. 727

6 Key Issues/Misconceptions with Thresholds 728

6.1 Which Thresholds Should be Used? 729

Unless there is clear reason to choose a different threshold 730
 value (e.g. political sensitivity), empirical estimates pro- 731
 vide a more appropriate value of the threshold than his- 732
 torical/heuristic thresholds, which are based on judgement. 733
 The key question is whether supply-side thresholds (which 734
 aim to represent the opportunity cost of investment to the 735
 system, given budget constraints) or demand-side thresh- 736
 olds (WTP estimates that aim to reflect the value that 737
 society places on a QALY) should be used [10]. A recent 738
 systematic review of WTP per QALY studies suggested 739
 that WTP per QALY varied substantially by condition, 740
 especially those for extending or saving life and improving 741
 QoL [73]. Supply-side thresholds enable the quantification 742
 of the net health gains (or losses) that would result from the 743
 inclusion of a new intervention (whether doing so repre- 744
 sents an increase in the budget or displaces a currently 745
 funded intervention[s]) in the healthcare system. Decisions 746
 made on the basis of supply-side CE thresholds ensure that 747
 aggregate health is improved by the inclusion of new 748
 interventions. 749

On the other hand, thresholds based on WTP for a 750
 QALY are generally higher than thresholds resulting from 751
 estimating the opportunity cost to the healthcare system 752
 [10]. As such, using WTP estimates may lead to decisions 753
 that reduce rather than improve health outcomes overall. 754
 This may also be the case with the use of WHO-CHOICE 755
 guidelines for thresholds (i.e. one to three times the GDP), 756
 where the threshold is not related to the efficiency of the 757
 healthcare system. However, as WTP estimates reflect 758
 societal WTP for improvements in health, the fact that they 759
 tend to be higher than estimates linked to the efficiency of 760
 the healthcare system provides suggestive evidence for an 761
 increase in public budgets for healthcare. Some analysts 762
 have argued that in a privately funded healthcare system, in 763

764 the absence of explicit healthcare budget constraint, WTP
765 can be an estimate of the opportunity cost of private con-
766 sumption [2].

767 6.2 Should the Threshold be Made Explicit?

768 There are two questions here: (1) whether there can be a
769 single threshold, and (2) whether the threshold values
770 should be made public. No HTA organization currently
771 recommends the use of a single threshold, and many do not
772 explicitly specify a threshold at all (as seen in Sect. 3).
773 Those that specify a threshold tend to specify a range rather
774 than a single value reflecting the belief that a single
775 threshold should not be applied to the diverse range of
776 technologies and conditions. In terms of the second ques-
777 tion, the so-called ‘silence of the lambda’ [74] or reluc-
778 tance to set out an explicit threshold, may result from a
779 number of concerns, including fear of gaming by phar-
780 maceutical companies to target ICERs just below the
781 threshold, reduced flexibility to balance competing criteria
782 when making funding decisions, and the issues associated
783 with advocating a threshold value that may have little or no
784 empirical basis (such as the potential for political and
785 ethical concerns about the accuracy and validity of funding
786 decisions) [75].

787 6.3 Impact of Using the Wrong Threshold

788 If the threshold used is lower than the empirical estimate, it
789 may lead to potentially cost-effective (compared with the
790 empirical threshold) technologies not being reimbursed.
791 However, it should be noted that in situations where
792 researchers suggested increasing the threshold [76], argu-
793 ments were based on WTP/preference estimates. On the
794 other hand, if the threshold used is higher than the
795 empirical estimate reflecting health opportunity costs, each
796 new technology approved (with a higher ICER than the
797 empirical threshold) leads to loss in health outcomes. An
798 example is NICE’s end-of-life decision-making scheme,
799 where it was suggested that approving drugs with an ICER
800 higher than the NICE threshold of £20 000–£30 000/
801 QALY resulted in substantial QALY losses [77]. Further-
802 more, Claxton et al. [31] argue that the current NICE
803 threshold (of £20 000–£30 000/QALY) is too high com-
804 pared with the empirical estimates, suggesting that
805 approving drugs lead to more health likely to be lost than
806 gained.

807 6.4 Threshold and Budget Impact

808 If the budget impact of a new technology is substantial (i.e.
809 non-marginal), the threshold used should be lower,
810 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 hep-
atitis 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 intro-
duction 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 poten-
tially 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 effec-
tive over time), the threshold will fall irrespective of
inflation. This argument is also applicable for the trans-
ferability 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 deter-
mined by estimating the efficiency of the healthcare sys-
tem, as observed in Sect. 4.

6.6 Threshold and Capacity Constraints

Published CEA studies often ignore the capacity con-
straints of resources (e.g. beds, nurses, equipment, etc.),
which may result in biased estimates of CE [82]. In prin-
ciple, 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 avail-
able 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

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

866 **6.7 Priority-Setting Process**

867 Alongside the results of CEA, a number of other factors are
 868 often also considered as part of the appraisal process
 869 around whether to adopt or reject an intervention. A recent
 870 review of all HTA appraisals between May 2000 to May
 871 2014 from the NICE, PBAC, Scottish Medicines Consor-
 872 tium, and CADTH suggested that technologies with ICERs
 873 higher than the respective thresholds are sometimes rec-
 874 ommended; the reasons included high clinical benefit over
 875 the standard of care, and addressing an unmet therapeutic
 876 need [84]. Similarly, even though some technologies (such
 877 as orphan drugs for rare diseases, or cancer treatments at
 878 end of life) have very high ICERs, NICE and most other
 879 health systems have found ways to fund those few tech-
 880 nologies on the basis of evidence of benefit. On the other
 881 hand, some interventions are rejected, even when the
 882 ICERs are below the threshold [46]. Indeed, it is
 883 acknowledged that there is a need for some discretion in
 884 priority setting linked to legitimization of decisions rather
 885 than using the threshold alone.

886 **7 Conclusions**

887 This paper contributes to the literature on CE thresholds by
 888 providing a simple illustration of the CE threshold as the
 889 shadow price of budget constraint, providing a theoretical
 890 framework for how a CE threshold could be employed in a
 891 hypothetical optimization setting. Existing estimates of
 892 ‘thresholds’ representing various definitions, from heuris-
 893 tics applied historically to more recent empirical estimates,
 894 whether WTP for improvements in health or opportunity
 895 costs are then outlined. Among these, those that can be
 896 categorized as supply-side estimates (i.e. from the UK,
 897 Australia, Spain and LMICs, as presented in Sect. 4.1) may
 898 be considered more appropriate for judging the CE of new
 899 technologies where the aim of agencies is to inform whe-
 900 ther or not a new technology is expected to improve pop-
 901 ulation health. Finally, the future for CE thresholds is
 902 speculated upon where new policy questions have indi-
 903 cated further areas of research where thresholds will be
 904 relevant and useful for decision making, particularly the
 905 consideration of effects and costs on multiple sectors
 906 beyond health where opportunity costs are still relevant.
 907 Despite advances in this area of research, there remain

misconceptions about CE thresholds, the assumptions 908
 involved and their implications, which this paper aimed to 909
 highlight. It is the responsibility of all of us to educate 910
 those who are involved in priority setting about these 911
 concepts of threshold in order to ensure efficient healthcare 912
 resource allocation. 913

Acknowledgements The authors would like to thank Brian Reddy 914
 and Simon Walker for their advice on the Irish threshold and eco- 915
 nomic evaluation across multiple sectors, respectively. 916

Author contributions PT conceived the initial idea for the paper, 917
 and all authors (PT, JO, AL, TT) contributed to developing the outline 918
 of the paper. TT led on the hypothetical example, including the 919
 development of the Microsoft Excel file in the ESM; JO led on the 920
 sections on empirical estimates of the supply-side threshold; AL led 921
 on the section about the WHO-CHOICE guidelines and contributed to 922
 the demand side thresholds; and PT led on the remaining sections in 923
 the paper. All authors were involved in the multiple revisions of the 924
 paper before signing off on the final version. 925

Funding This research received no specific grant from any funding 926
 agency in the public, commercial, or not-for-profit sectors. 927

Compliance with ethical standards 928

Conflict of interest Praveen Thokala, Jessica Ochalek, Ashley 929
 Leech, and Thaison Tong have no conflicts of interest to declare. 930
 931

References 932

1. Neumann PJ, Thorat T, Zhong Y, Anderson J, Farquhar M, Salem 933
 M, et al. A systematic review of cost-effectiveness studies reporting 934
 cost-per-DALY averted. *PloS One*. 2016;11(12):e0168512. 935

2. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance 936
 GW. *Methods for the economic evaluation of health care pro-* 937
grammes. 4th ed. Oxford: Oxford University Press; 2015. p. 464. 938

3. *Guide to the methods of technology appraisal 2013. Guidance* 939
and guidelines. NICE. [https://www.nice.org.uk/process/pmg9/](https://www.nice.org.uk/process/pmg9/chapter/foreword) 940
[chapter/foreword](https://www.nice.org.uk/process/pmg9/chapter/foreword). Accessed 14 Dec 2017. 941

4. *Guidelines for the Economic Evaluation of Health Technologies:* 942
Canada. CADTH.ca [https://www.cadth.ca/about-cadth/how-we-](https://www.cadth.ca/about-cadth/how-we-do-it/methods-and-guidelines/guidelines-for-the-economic-evaluation-of-health-technologies-canada) 943
[do-it/methods-and-guidelines/guidelines-for-the-economic-evalu-](https://www.cadth.ca/about-cadth/how-we-do-it/methods-and-guidelines/guidelines-for-the-economic-evaluation-of-health-technologies-canada) 944
[ation-of-health-technologies-canada](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. 945

5. Australian Government Department of Health. *Guidelines for* 946
preparing a submission to the Pharmaceutical Benefits Advisory 947
Committee: Version 5.0. 2016. [https://pbac.pbs.gov.au/content/](https://pbac.pbs.gov.au/content/information/files/pbac-guidelines-version-5.pdf) 948
[information/files/pbac-guidelines-version-5.pdf](https://pbac.pbs.gov.au/content/information/files/pbac-guidelines-version-5.pdf). Accessed Dec 949
 2017. 950

6. Sculpher M, Claxton K, Pearson SD. Developing a value 951
 framework: the need to reflect the opportunity costs of funding 952
 decisions. *Value Health*. 2017;20(2):234–9. 953

7. Culyer T. *Cost-effectiveness thresholds in health care: a book-* 954
shelf guide to their meaning and use. York: Centre for Health 955
 Economics, University of York; 2015. p. 1–22. 956

8. Bertram MY, Lauer JA, De Joncheere K, Edejer T, Hutubessy R, 957
 Kieny M-P, et al. Cost-effectiveness thresholds: pros and cons. 958
Bull World Health Organ. 2016;94(12):925–30. 959

9. Woods B, Revill P, Sculpher M, Claxton K. Country-level cost- 960
 effectiveness thresholds: initial estimates and the need for further 961
 research. *Value Health*. 2016;19(8):929–35. 962

- 963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
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*. *Macroeconomie 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, Hammit 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-middle-income-countries\(12487fa5-e63f-4ac3-9fa4-03b2795065eb\).html](https://pure.york.ac.uk/portal/en/publications/cost-per-daly-averted-thresholds-for-low-and-middle-income-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).

1095 [https://www.oecd.org/health/Value-in-Pharmaceutical-Pricing-](https://www.oecd.org/health/Value-in-Pharmaceutical-Pricing-Australia.pdf)
 1096 [Australia.pdf](https://www.oecd.org/health/Value-in-Pharmaceutical-Pricing-Australia.pdf). Accessed Dec 2017.

1097 45. Henry DA, Hill SR, Harris A. Drug prices and value for money:
 1098 the Australian pharmaceutical benefits scheme. *JAMA*.
 1099 2005;294(20):2630–2.

1100 46. Griffiths EA, Vadlamudi NK. Cadth’s \$50,000 cost-effectiveness
 1101 threshold: fact or fiction? 2016; ISPOR. [https://www.ispor.org/](https://www.ispor.org/research_pdfs/54/pdffiles/PHP278.pdf)
 1102 [research_pdfs/54/pdffiles/PHP278.pdf](https://www.ispor.org/research_pdfs/54/pdffiles/PHP278.pdf). Accessed Dec 2017.

1103 47. Grocott R, Metcalfe S, Alexander P, Werner R. Assessing the
 1104 value for money of pharmaceuticals in New Zealand: PHAR-
 1105 MAC’s approach to cost-utility analysis. *N Z Med J*.
 1106 2013;126(1378):60–73.

1107 48. O’Keeffe KM, Gander PH, Scott WG, Scott HM. Insomnia
 1108 treatment in New Zealand. *N Z Med J*. 2012;125(1349):46–59.

1109 49. Simoens S. Health economic assessment: a methodological primer.
 1110 *Int J Environ Res Public Health*. 2009;6(12):2950–66.

1111 50. Metcalfe S, Rodgers A, Werner R, Schousboe C. PHARMAC has
 1112 no cost-effectiveness threshold. *N Z Med J*.
 1113 2012;125(1350):99–101.

1114 51. Metcalfe S, Grocott R. Comments on “Simoens, S. Health economic
 1115 assessment: a methodological primer. *Int. J. Environ. Res. Public Health* 2009, 6, 2950–2966”—New Zealand in fact has no
 1116 cost-effectiveness threshold. *Int J Environ Res Public Health*.
 1117 2010;7(4):1831–4.

1118 52. Scottish Medicines Consortium. SMC Modifiers used in
 1119 Appraising New Medicines. [https://www.scottishmedicines.org.](https://www.scottishmedicines.org.uk/About_SMC/Policy_statements/SMC_Modifiers_used_in_Appraising_New_Medicines)
 1120 [uk/About_SMC/Policy_statements/SMC_Modifiers_used_in_](https://www.scottishmedicines.org.uk/About_SMC/Policy_statements/SMC_Modifiers_used_in_Appraising_New_Medicines)
 1121 [Appraising_New_Medicines](https://www.scottishmedicines.org.uk/About_SMC/Policy_statements/SMC_Modifiers_used_in_Appraising_New_Medicines). Accessed 14 Dec 2017.

1122 53. Kamae I. Value-based approaches to healthcare systems and
 1123 pharmacoconomics requirements in Asia: South Korea, Taiwan,
 1124 Thailand and Japan. *Pharmacoconomics*. 2010;28(10):831–8.

1125 54. Coelho De Soarez P, Dutilh Novaes HM. Cost-effectiveness
 1126 thresholds and the Brazilian Unified National Health System.
 1127 2017. [http://www.scielo.br/pdf/csp/v33n4/1678-4464-csp-33-04-](http://www.scielo.br/pdf/csp/v33n4/1678-4464-csp-33-04-e00040717.pdf)
 1128 [e00040717.pdf](http://www.scielo.br/pdf/csp/v33n4/1678-4464-csp-33-04-e00040717.pdf). Accessed Dec 2017.

1129 55. Gravelle HS, Backhouse ME. International cross-section analysis
 1130 of the determination of mortality. *Soc Sci Med*.
 1131 1987;25(5):427–41.

1132 56. Barnsley P, Towse A, Karlsberg Schaffer S, Sussex J. Critique of
 1133 CHE Research Paper 81: methods for the estimation of the NICE
 1134 cost effectiveness threshold. OHE. 2013. [https://www.ohe.org/](https://www.ohe.org/publications/critique-che-research-paper-81-methods-estimation-nice-cost-effectiveness-threshold)
 1135 [publications/critique-che-research-paper-81-methods-estimation-](https://www.ohe.org/publications/critique-che-research-paper-81-methods-estimation-nice-cost-effectiveness-threshold)
 1136 [nice-cost-effectiveness-threshold](https://www.ohe.org/publications/critique-che-research-paper-81-methods-estimation-nice-cost-effectiveness-threshold). Accessed 14 Dec 2017.

1137 57. Claxton K, Sculpher M. Response to the OHE critique of CHE
 1138 Research paper 81. [https://www.york.ac.uk/media/che/](https://www.york.ac.uk/media/che/documents/Response%20to%20the%20OHE%20critique%20of%20CHE%20Research%20paper%2081.pdf)
 1139 [documents/Response%20to%20the%20OHE%20critique%](https://www.york.ac.uk/media/che/documents/Response%20to%20the%20OHE%20critique%20of%20CHE%20Research%20paper%2081.pdf)
 1140 [20of%20CHE%20Research%20paper%2081.pdf](https://www.york.ac.uk/media/che/documents/Response%20to%20the%20OHE%20critique%20of%20CHE%20Research%20paper%2081.pdf). Accessed Dec
 1141 2017.

1142 58. Edney L, Afzali H, Cheng T, Karnon J. Estimating the reference
 1143 ICER for the Australian health system. *Pharmacoconomics*.
 1144 2017. <https://doi.org/10.1007/s40273-017-0585-2> (Epub 22 Dec
 1145 2017).

1146 59. Vallejo-Torres L, García-Lorenzo B, Serrano-Aguilar P. Esti-
 1147 mating a cost-effectiveness threshold for the Spanish NHS.
 1148 FEDEA; 2016 Jun. Report No. eee2016-22. [https://www.ucl.ac.](https://www.ucl.ac.uk/dahr/pdf/HESG/Paper_A31.pdf)
 1149 [uk/dahr/pdf/HESG/Paper_A31.pdf](https://www.ucl.ac.uk/dahr/pdf/HESG/Paper_A31.pdf). Accessed Dec 2017.

1150 60. Government of the Republic of Malawi. Health Sector Strategic
 1151 Plan II (2017–2022): towards Universal Coverage. 2017. file:///Users/ashleylech/Downloads/HSSP%20II%20Final%20HQ%20complete%20file.pdf.pdf. Accessed Dec 2017.

1152 61. Ochalek J, Claxton K, Revill P, Sculpher M, Rollinger A. Sup-
 1153 porting the development of an essential health package: princi-
 1154 ples and initial assessment for Malawi. Centre for Health
 1155 Economics, University of York; 2016 Sep. (Working Papers).
 1156 Report No.: 136cherp [https://www.york.ac.uk/media/che/](https://www.york.ac.uk/media/che/documents/papers/researchpapers/CHERP136_EHP_Malawi_interventions.pdf)
 1157 [documents/papers/researchpapers/CHERP136_EHP_Malawi_](https://www.york.ac.uk/media/che/documents/papers/researchpapers/CHERP136_EHP_Malawi_interventions.pdf)
 1158 [interventions.pdf](https://www.york.ac.uk/media/che/documents/papers/researchpapers/CHERP136_EHP_Malawi_interventions.pdf). Accessed 17 Oct 2017.

1159 62. Thavorncharoensap M, Teerawattananon Y, Natanant S, Kulpeng
 W, Yothasamut J, Weraingyong 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. Determina-
 tion of cost-effectiveness threshold for health care interven-
 tions in Malaysia. *Value Health*. 2017;20(8):1131–8.

65. GDP per capita (current US\$). Data. [https://data.worldbank.org/](https://data.worldbank.org/indicator/NY.GDP.PCAP.CD)
[indicator/NY.GDP.PCAP.CD](https://data.worldbank.org/indicator/NY.GDP.PCAP.CD). Accessed 14 Dec 2017.

66. Initiative on US Value Assessment Frameworks. [https://www.](https://www.ispor.org/ValueAssessmentFrameworks/Index)
[ispor.org/ValueAssessmentFrameworks/Index](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 prac-
 tices, 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/](https://healthpolicy.fsi.stanford.edu/sites/default/files/ohe_hta_methods.pdf)
[files/ohe_hta_methods.pdf](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 deci-
 sion 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? *Pharmacoconomics*. 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-effective-
 ness 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/](https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/)
[About/what-we-do/NICE-guidance/NICE-technology-appraisals/](https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/)

- 1226 [TA-HST-procedure-varying-the-funding-direction.pdf](#). Accessed
1227 14 Dec 2017. 1234
- 1228 81. Ubel PA, Hirth RA, Chernew ME, Fendrick AM. What is the
1229 price of life and why doesn't it increase at the rate of inflation? 1235
1230 Arch Intern Med. 2003;163(14):1637–41. 1236
- 1231 82. Thokala P, Dixon S, Jahn B. Resource modelling: the missing
1232 piece of the HTA jigsaw? Pharmacoeconomics. 1237
1233 2015;33(3):193–203. 1238
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. 1239
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. 1240
1241
1242

REVISED PROOF