


Revealed and Stated Preferences of Decision Makers for Priority Setting in Health Technology Assessment: A Systematic Review

Peter Ghijben¹  · Yuanyuan Gu² · Emily Lancsar¹ · Silva Zavarsek^{1,3}

Published online: 9 November 2017
© Springer International Publishing AG 2017

Abstract

Background There is much interest from stakeholders in understanding how health technology assessment (HTA) committees make national funding decisions for health technologies. A growing literature has analysed past decisions by committees (revealed preference, RP studies) and hypothetical decisions by committee members (stated preference, SP studies) to identify factors influencing decisions and assess their importance.

Objectives A systematic review of the literature was undertaken to provide insight into committee preferences for these factors (after controlling for other factors) and the methods used to elicit them.

Methods Ovid Medline, Embase, Econlit and Web of Science were searched from inception to 11 May 2017. Included studies had to have investigated factors considered by HTA committees and to have conducted multivariate analysis to identify the effect of each factor on funding decisions. Factors were classified as being important based on statistical significance, and their impact on decisions was compared using marginal effects.

Results Twenty-three RP and four SP studies (containing 42 analyses) of 14 HTA committees met the inclusion criteria. Although factors were defined differently, the SP literature generally found clinical efficacy, cost-effectiveness and equity factors (such as disease severity) were each important to the Pharmaceutical Benefits Advisory Committee (PBAC), the National Institute for Health and Care Excellence (NICE) and the All Wales Medicines Strategy Group. These findings were supported by the RP studies of the PBAC, but not the other committees, which found funding decisions by these and other committees were mostly influenced by the acceptance of the clinical evidence and, where applicable, cost-effectiveness. Trust in the evidence was very important for decision makers, equivalent to reducing the incremental cost-effectiveness ratio (cost per quality-adjusted life-year) by A\$38,000 (Australian dollars) for the PBAC and £15,000 for NICE.

Conclusions This review found trust in the clinical evidence and, where applicable, cost-effectiveness were important for decision makers. Many methodological differences likely contributed to the diversity in some of the other findings across studies of the same committee. Further work is needed to better understand how competing factors are valued by different HTA committees.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s40273-017-0586-1>) contains supplementary material, which is available to authorized users.

✉ Peter Ghijben
peter.ghijben@monash.edu

¹ Centre for Health Economics, Monash Business School, Monash University, Clayton, VIC 3800, Australia

² Centre for the Health Economy, Macquarie University, North Ryde, NSW 2109, Australia

³ Deakin Health Economics, Centre for Population Health Research, Deakin University, Geelong, VIC 3220, Australia

Key Points

Stated preference (SP) studies found clinical efficacy, cost-effectiveness and equity concerns were all important factors in decisions made by the Pharmaceutical Benefits Advisory Committee (PBAC), National Institute for Health and Clinical Excellence (NICE) and All Wales Medicines Strategy Group.

Depending on the committee, revealed preference (RP) studies typically found funding decisions were driven by trust in the clinical evidence and, where applicable, cost-effectiveness; the role of equity factors based on past decisions was unclear.

The impact of these key factors (trust in the clinical evidence and cost-effectiveness) on decisions was similar across RP studies of the PBAC and NICE, despite methodological differences across the literature.

1 Introduction

Decisions to fund new and expensive health technologies on public insurance schemes compel decision makers to balance the demand for services with the reality of constrained budgets. Ultimately, this requires decision makers to set priorities for the provision of health care and trade-off between competing objectives (such as health gains, equity considerations and cost). Over the past 25 years, many countries have established agencies and committees to conduct health technology assessment (HTA) and make funding decisions [1]. A key ethical consideration is that these decisions must be socially justifiable, which has prompted much debate regarding which factors should be considered and to what extent [2, 3]. Unsurprisingly there is also considerable interest from the public and other stakeholders in understanding which factors actually influence decisions. A large number of studies have been published over the past two decades investigating HTA committees' decisions, to understand how they are made [4–8]. This broad literature has addressed three key research questions: (1) What processes are followed for decision making? (2) Which factors are important? (3) How important are those factors?

Early studies typically adopted qualitative methods, such as observing or interviewing decision makers, and mostly addressed the first two research questions. Reviews

based largely on these studies found (1) HTA committees use an explicit or implicit set of criteria, but decisions are inherently 'value-based' [4]; (2) similar factors (broadly defined including clinical efficacy, cost-effectiveness, affordability, and need) are considered by different committees, but there was little or no information on how any criteria were defined or used [5]; and (3) clinical evidence was generally assessed before cost-effectiveness and other considerations [6, 7].

In contrast, quantitative studies are more recent and have focused on the second and third questions by analysing past decisions (revealed preference, RP studies) or hypothetical decisions (stated preference, SP studies) to elicit preferences from decision makers. To our knowledge, only one review has focused on the quantitative literature which found that a large and diverse number of factors have been explored, related mainly to technology characteristics and appraisal methods used by committees [8]. However, that review, which was mostly descriptive, did not systematically identify which factors were demonstrated to be important or quantify their impact on decisions; nor did it include SP studies. Therefore, relatively little is known about what the findings of the quantitative literature are as a whole as to which factors influence funding decisions and to what extent.

This systematic review was undertaken to address these important unanswered questions from the quantitative literature. The methods used by studies were also examined because understanding how their results were generated is useful in identifying potential limitations of this literature. Specifically, this review addresses the following questions: (1) What methods have been used to elicit revealed and stated preferences from HTA decision makers for evidence considered in funding decisions? (2) Which factors have been found to be important in determining decisions after controlling for other factors? (3) What is the impact of those factors on decisions? (4) What differences exist between RP and SP studies?

2 Methods

2.1 Overview

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) reporting guidelines [9]. To identify the effect of factors after controlling for the influence of other variables, only studies which presented

multivariate analysis of past or hypothetical funding decisions by HTA decision makers were included.¹

2.2 Literature Search

Four databases were searched: Ovid Medline (1946 to present with daily update), Embase, Econlit and Web of Science [Science Citation Index (SCI)-expanded and Social Sciences Citation Index (SSCI)] from conception to 11 May 2017. Search terms were developed for two categories: (1) health care funding decisions and (2) preferences. The search terms were developed for the Ovid Medline database and modified for the respective databases. The complete search histories are provided in the electronic supplementary material. Backward searching of references and forward searching of citations from included studies were conducted to identify additional studies for screening ('hand search').

2.3 Eligibility Criteria

Inclusion and exclusion criteria are summarised below, with more detail provided in the electronic supplementary material. Studies not published in a peer reviewed publication or in English were excluded. Studies were included if they (1) investigated factors considered by committees and their impact on recommendations for approving health technologies on publically funded national insurance schemes; and (2) conducted multivariate analysis (to identify the influence of each factor on the decision after controlling for other factors). For SP studies, preferences had to have been elicited from HTA decision makers (past or current members of either an HTA committee or sub-committee) between competing criteria in the context of funding decisions. Samples of other experts or the general public were excluded, including mixed samples with decision makers where the preferences of committee members could not be separately identified. As the review focused on committee preferences from evidence considered, key variables defined in RP studies had to have been based on data from official committee documents. For example, studies which sourced incremental cost-effectiveness ratios (ICERs) for given interventions from the international literature, rather than from committee documents, were excluded. It was considered that given ICERs are dependent on the context (such as population/

indication, comparator and costs) as well as other assumptions, ICERs published in the literature may not be relevant to specific committees and their decisions. Two authors (PG and YG) independently screened the titles and abstracts of all the studies identified from the search strategies. The inclusion/exclusion criteria were then sequentially applied to the full text of all retrieved citations. Discrepancies were resolved by discussion, until a consensus between the authors was reached.

2.4 Data Extraction and Synthesis

Data were extracted from the included studies by two authors (PG and SZ). Multiple analyses from each study were included if they were based on a different (1) sample; (2) econometric model; (3) model specification; or (4) key variable definition. However, where stepwise procedures were used and multiple stepped analyses presented, only results from the preferred model were included. Sensitivity analyses and variables (including redefined variables) only presented in supplementary materials were excluded given incomplete reporting across studies. For the synthesis and interpretation of factors, all variables defined and modelled by the studies were categorised on the basis of taxonomies in the literature [10–12] as either (1) content of the decision ('content'); (2) processes of the committee ('process'); or (3) 'other'. Content variables were subcategorised as relating to (1) therapeutic value; (2) economic value; (3) disease characteristics; and (4) technology characteristics. Variables were classified as being important on the basis of statistical significance ($p \leq 0.05$) for parametric models, or where included in the final 'pruned' tree in classification and regression tree (CART) analysis. To explore the impact of factors on decisions across studies, marginal effects from a fixed reference point were calculated [13]. Coefficients from parametric models were used to calculate the change in the probability of a positive recommendation from a reference of 50% for a given change in each factor.² This analysis was performed for two committees for which sufficient data were available (the Pharmaceutical Benefits Advisory Committee or PBAC, and the appraisal committee of the National Institute for Health and Clinical Excellence or NICE) and three factors (ICER, budget impact and issue with the clinical evidence) which were commonly defined across studies. Where ICER and budget impact were modelled as categorical variables, the marginal effect was calculated based on the mean values of the categories. For comparison across countries, currency was converted using purchasing power parities and budget impact was adjusted

¹ Univariate analyses were excluded because they do not account for confounding effects of other factors which may lead to biased results. For example, given cancer drugs are generally more expensive than non-cancer drugs and more expensive drugs are less likely to be funded than less expensive drugs, a univariate analysis may find that cancer drugs are less likely to be funded than non-cancer drugs. However, this may not be the case if the analysis also controls for drug costs.

² For example, in a logistic model without interaction terms, the change in probability of a positive recommendation from a reference point $\Delta P_j = \frac{1}{1+e^{-(L+\beta_j)}} - P_{\text{ref}}$, where $L = \ln\left(\frac{P_{\text{ref}}}{1-P_{\text{ref}}}\right)$.

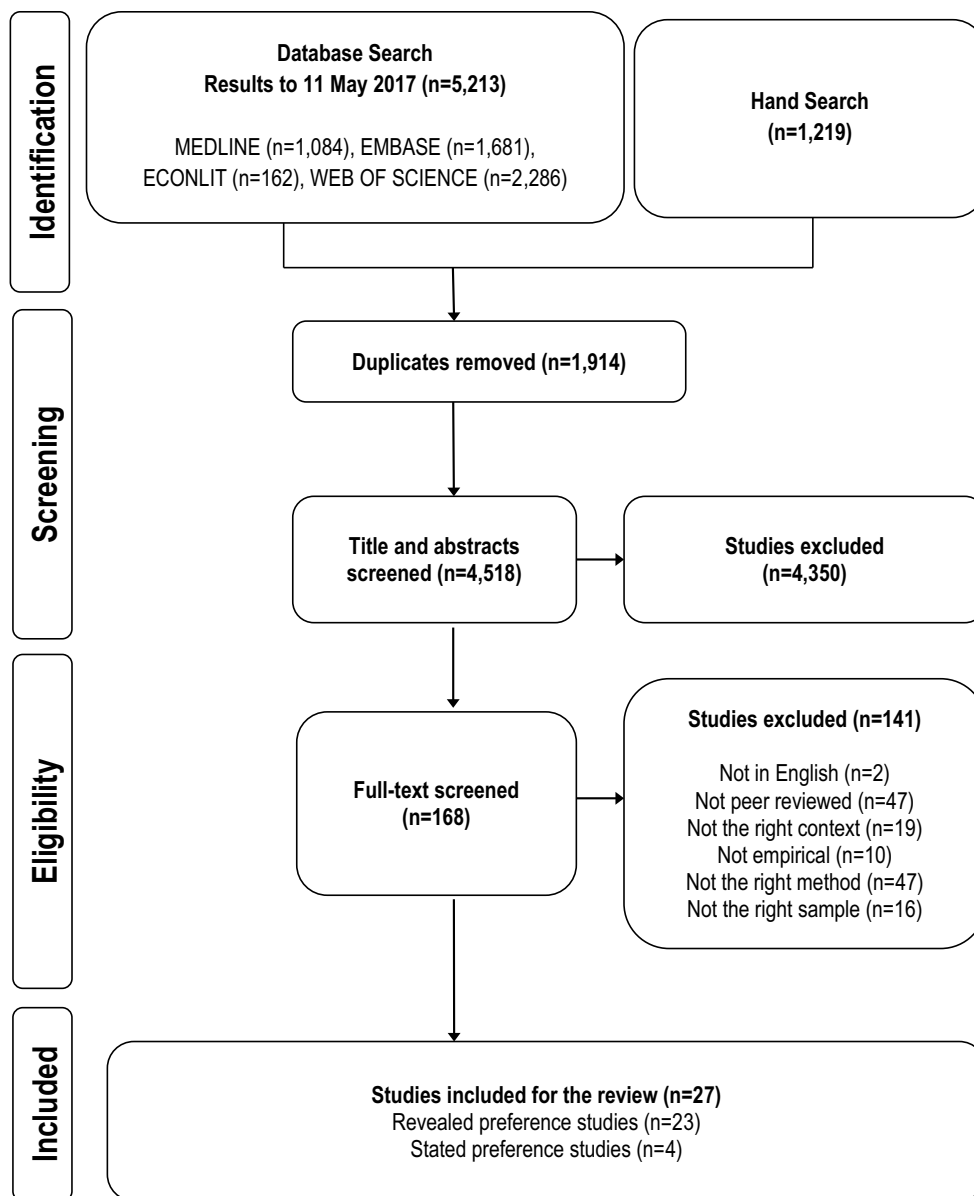


Fig. 1 Flow diagram for selection of studies

to account for the relative size of government health budgets [14]. Meta-regression was not performed given extensive overlap in the HTA decisions modelled across studies (see the electronic supplementary material) [15].

3 Results

3.1 Overview

The selection of studies is summarised in Fig. 1. The searches identified 4518 unique citations for screening. Full-text articles were retrieved for 168 references, and 27 studies were included in the review. Several studies

[16–22] which presented multivariate analysis were excluded because they did not elicit preferences from HTA decision makers or use evidence considered by them.³ The

³ A number of these used ICERs sourced from the literature rather than ICERs considered by the committees. However, none of these studies did so due to confidentiality issues. These studies were exploratory in nature and did not investigate the impact of evidence considered on funding decisions. In one study [18], the committee under consideration does not assess cost-effectiveness and hence no ‘official’ ICER existed. In the other studies [17, 19], ‘current funding status’ of health technologies was investigated. However, ‘current funding status’ is not equivalent to a past funding decision, given technologies which are not currently funded may never have been considered by a committee. None of the data used in those studies were sourced from committee documents.

complete list of studies and their methods are summarised in the electronic supplementary material. For clarity, decision makers will be referred to in this review by committee or agency name, or more generally as 'committee'.

The majority of studies (23 of 27) included in the review were RP studies [23–45], although four SP studies [46–49] met the inclusion criteria. All studies investigated preferences from a single committee, with the exceptions of Cerri et al. [45], which pooled past recommendations from four committees (but only modelled a fixed effect per committee), and Kim et al. [41], which analysed final government decisions (rather than committee recommendations that may differ). Although the findings from both studies could not be attributed to individual committees, their methods were still reviewed. Excluding Cerri et al. [45] and Kim et al. [41], the most commonly studied committees were the PBAC in Australia ($n = 6$), NICE in England and Wales ($n = 6$), and the All Wales Medicines Strategy Group (AWMSG) in Wales ($n = 2$). Members from these three committees (including sub-committees) were each the subject of one SP study. The remaining 11 studies (ten RP and one SP) investigated preferences of 11 different committees, presented in Fig. 2.

Overall, most studies analysed funding decisions for pharmaceuticals, which is unsurprising given the majority of committees included in the review (11 of 14) are only responsible for appraising pharmaceuticals. Of the eight studies on committees which also evaluate other types of health technologies, decisions for non-pharmaceuticals were excluded by three (see electronic supplementary material for details).

Data were extracted from 36 RP and six SP analyses across the 27 included studies. The median number of past HTA decisions analysed was 138 (range 33–1453), whereas the SP studies enrolled between 11 and 41 decision makers, who collectively made between 218 and 1148 hypothetical decisions. The median number of explanatory variables defined and modelled in the RP analyses was 15.5 (range 4–50) and seven (range 2–20), respectively. Two RP studies [23, 26] modelled at least one interaction between two or more variables, although they were defined but not modelled by three others [24, 32, 44]. In contrast, a median of seven (range 5–12) variables were defined and modelled without interactions in the SP studies.

The majority of explanatory variables ($\sim 85\%$) defined across the RP studies were classified as content factors associated with some aspect of therapeutic value, economic value or disease/technology characteristics. Committee-specific process factors such as the participation of patient advocates with NICE or AWMSG and prior consideration by the PBAC were also commonly defined across the literature. A wide range of other factors, including date of

decision, committee composition, national statistics or elections, publicity and revenue of the manufacturer, were less commonly explored. In contrast, the SP literature focused exclusively on the trade-off between clinical, economic and equity factors, all else assumed constant.

3.2 Methods for Eliciting Revealed Preferences

3.2.1 Data Collection

Data used in the majority of RP studies (21 of 23) were generated with researcher-driven protocols to collect and define variables from committee documents. Commercial databases of HTA decisions and appraisal criteria (based on committee documents) were used by two studies [32, 36]. Access to complete and reliable data was a common limitation across the literature. It precluded three studies from modelling ICER despite playing a central role in decision making [34, 40, 41] and prompted differences in how ICER was defined by others. For example, where the 'official' ICER considered by NICE was not available, one study [28] used the mean of the reported ICERs, two studies [29, 31] used the proposed (unappraised) ICER or mean, and one study [32] used simulation to predict it. Only four studies used data from confidential sources [23, 26, 37, 38].

3.2.2 Sample

Selection criteria were applied by studies to identify samples of decisions for analysis. The most common criteria related to the type of economic analysis and the ICER. For example, nine studies [23, 26, 28–30, 32, 33, 42, 44] only included decisions based on a cost-effectiveness analysis (CEA) with an ICER (in at least one analysis) expressed only as cost per quality-adjusted life-year (QALY) [23, 26, 30, 32, 33, 42, 44] or including cost per life-year (LY) [28, 29]. Three of the nine studies [23, 26, 32] then excluded decisions where the technology was either more costly but less effective or less costly but more effective; a fourth study [42] excluded the former but not the latter to preserve sample size. The rationale was that excluded decisions based on cost-minimisation analysis (CMA) or other well-defined decision criteria do not reflect a trade-off between health gains, equity considerations and cost. One exception was the study of Linley and Hughes [33], which included decisions from CMAs as if they were based on CEA by assuming the implied ICER is zero. Three of these nine studies [32, 42, 44] also excluded decisions for therapies both less costly and less effective, because the loss of health may be valued differently by committees.

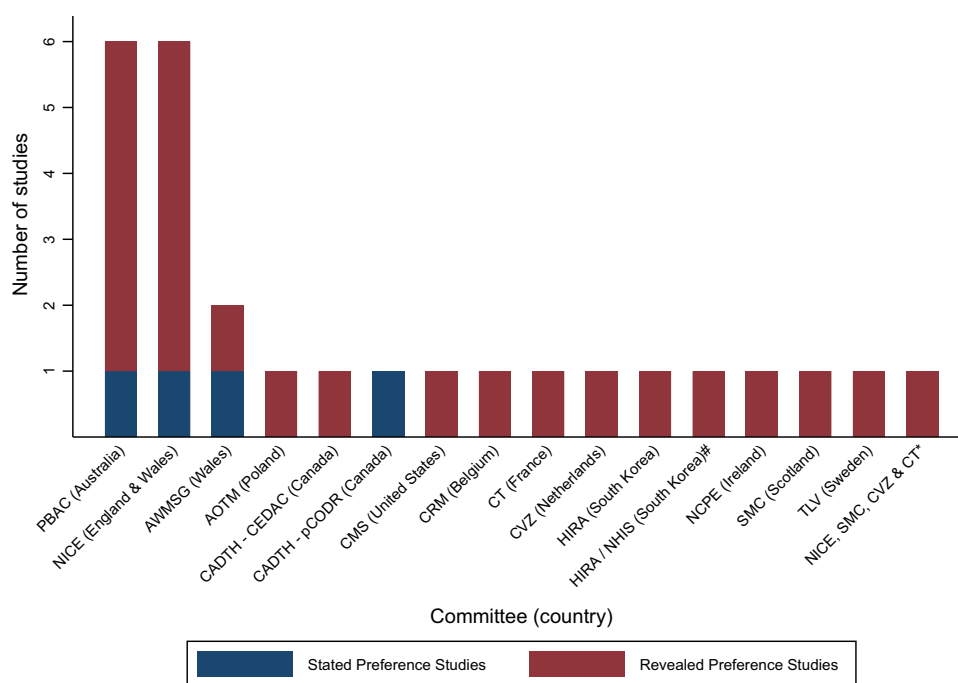


Fig. 2 Distribution of studies by committee and preference type. *AOTM* Agencja Oceny Technologii Medycznych (Agency for Health Technology Assessment), *AWMSG* All Wales Medicines Strategy Group, *CADTH* Canadian Agency for Drugs and Technologies in Health, *CEDAC* Canadian Expert Drug Advisory Committee, *CMS* Centres for Medicare and Medicaid Services, *CRM* Commission de Remboursement des Médicaments (The Commission of Reimbursement of Medicines), *CT* Commission de la Transparence (Transparency Commission), *CVZ* College Voor Zorgverzekering (Health Care Insurance Board), *HIRA* Health Insurance Review and

Assessment service, *NCPE* National Centre for Pharmacoeconomics, *NHIS* National Health Insurance Service, *NICE* National Institute for Health and Care Excellence, *PBAC* Pharmaceutical Benefits Advisory Committee, *pCODR* pan-Canadian Oncology Drug Review, *SMC* Scottish Medicines Consortium, *TLV* Tandvårds-och läkemedelsförmånsverket (Dental and Pharmaceutical Benefits Agency). # Kim et al. [41] analysed final government decision rather than committee recommendation. * Cerri et al. [45] pooled decisions across committees

Other selection criteria were mostly used to investigate specific diseases [27, 37, 38] or exclude special cases such as paediatric populations [31, 39, 45], ‘rule of rescue’ [40, 41] and conditional recommendations [37]. Three studies [25, 30, 35] in at least one of their analyses excluded previous or subsequent decisions related to the same therapy to avoid the correlation between related decisions. Otherwise all decisions were assumed to be unrelated or this was considered by using clustering [23, 26, 32], control variables [23–25, 27, 35, 43], robust standard errors [28, 29] or a generalised estimating equation approach [27]. Missing data were excluded by six studies [23–26, 30, 33], coded as categories (‘not stated’/‘not identified’/‘not reported’) by two [34, 35] and imputed (assuming missing at random) by four [31, 39, 42, 45].

3.2.3 Definition of Explanatory Variables

Although similar types of factors were explored across the studies, definitions varied greatly due in part to inconsistent reporting of data in official documents and use of ex-post

definitions.⁴ For example, nine studies defined six different variables related to severity of disease: (1) disability-adjusted life-year weight [32]; (2) greater impact on survival relative to quality of life [33]; (3) premature mortality (<5 years) [23, 26]; (4) ‘severe’ disease or ‘short’ duration of life [28, 42]; (5) ‘life-threatening’ disease [35]; and (6) a committee-assessed scale [38, 44]. Although most studies defined ICER in terms of cost per QALY, the inclusion of decisions without ICERs made interpretation problematic. For example, four studies [24, 25, 33, 43] assumed the ICERs for such decisions were zero or modelled them as separate categories, two studies [31, 45] imputed missing ICERs (despite not being relevant to all decisions), and four studies [27, 35, 37, 39] dropped the variable all together. One study [33] used ICER to define incremental

⁴ Variables were commonly defined on the basis of an interpretation of the source data by researchers rather than an explicit classification by the committee. Such variables were considered to be defined in an ex-post fashion, or after the fact, and may not have been directly considered by the committee.

net monetary benefit (INMB) at a given cost-effectiveness threshold (£20,000/QALY).

Factors associated with clinical uncertainty were also variably defined across studies and incompletely modelled with respect to the broader concept (see Table 1 and the electronic supplementary material for details). For example, quantity of evidence (by type of evidence) was defined by nine studies in terms of either number of studies, number of patients enrolled and/or duration of trials. However, none of the studies took account of whether the trials were powered to support the findings of the relevant outcome(s). Clinical uncertainty is also a function of level or hierarchy of evidence, direct or indirect evidence, quality or risk of bias, and relevance of evidence for the decision. Most studies (18 of 23) explored at least one of these factors to some extent with or without controlling for the others. Conversely, ten studies defined (and eight modelled) a variable associated with the committee's judgement of the clinical evidence (after consideration of these factors) and whether the evidence was accepted or not, either in addition to or instead of the separate variables.

3.2.4 Outcome Variable and Econometric Methods

Committee decisions were most commonly defined as either recommended or not recommended. Depending on the study and committee, a recommended but restricted category (restricted in terms of subsidised listing or population compared to the registered indication) was also common, but modelled differently. For example, it was defined either as a unique category [29, 31, 45], as part of an ordered variable by level of restriction [30], or as two unique decisions (not recommended for the unrestricted indication and recommended for the restricted indication) [28, 32] in the studies of NICE. Where applicable, other studies mostly pooled restricted and recommended decisions. This was justified on the basis of sample size [43] or the argument that not all committees make true restricted decisions where a narrower indication is recommended compared to the indication actually requested by the manufacturer [33]. Other studies also found consistent results for some committees when restricted decisions were redefined as recommended [39, 45]. Deferred (i.e. no decision) and conditional (i.e. temporal or coverage with evidence) decisions, where applicable and not excluded, were typically defined as not recommended, with two exceptions [39, 45]. In contrast to the other committees, the *Commission de la Transparence* (CT) makes recommendations in terms of absolute clinical benefit and, since 2004, incremental benefit, using graded scales [50]. One study modelled the former as an ordered variable [38], whereas another split the latter into three unique categories of incremental improvement [45].

Subsequently, past decisions were most commonly modelled using binary logit (16 studies, including 21 analyses) or probit (one study and two analyses) models. Others used a multinomial logit (MNL) model (four studies and six analyses), an ordered logit model (two studies and four analyses), a discrete time logit model (one study and one analysis), and CART (two studies and two analyses), where two studies [25, 27] used more than one class of model (see the electronic supplementary material for details).

3.2.5 Model Specification

The selection of explanatory variables in the multivariate analyses differed considerably across the literature. Overall, four studies [26, 28, 29, 44] (five analyses) selected variables predominately on theoretical grounds with or without model testing; 14 studies [24, 25, 27, 30, 31, 34–37, 39, 40, 42, 43, 45] (21 analyses) selected variables on the basis of statistical methods; and five studies [23, 32, 33, 38, 41] (ten analyses) used both approaches, for different analyses. The statistical methods employed also varied. For example, variables were included based on statistical significance in univariate analysis (seven analyses), the application of various stepwise procedures (backward elimination, forward selection or bi-directional methods) to an initial set of variables (14 analyses), or the Bolasso method [51] (three analyses). Criteria used for inclusion of variables in both CART models were not clear (see the electronic supplementary material for details).

3.3 Methods for Eliciting Stated Preferences

3.3.1 Data Generation and Sample

SP studies generated data by asking committee members to make hypothetical decisions. The four studies all used discrete-choice experiments (DCEs) to elicit preferences; however, one study combined these results with a different type of choice task [49]. All studies asked committee members to make repeated choices between different health technologies described by a set of variables. While a similar elicitation method was used by all studies, the experimental and survey designs differed markedly. Studies used either orthogonal [46, 47] or efficient [48, 49] experimental designs to generate the hypothetical alternatives and questions. Alternatives were described by five to eight unique factors defined across two to four levels. Choice sets consisted of either one [47] or two [46, 48, 49] hypothetical unlabelled alternatives described in absolute [46, 49] or relative [47, 48] values (to current therapy) depending on the wording of the questions. Unlike the RP

studies, all SP studies assumed new technologies were more costly but more effective. Survey response rates ranged from 29.7% to 56.9% (where reported) from a relatively small number of past or present committee or sub-committee members (see the electronic supplementary material for details).

3.3.2 Definition of Explanatory Variables

Although a similar set of efficacy, economic and equity factors were explored across studies, there were a number of differences in how variables were defined and modelled. Similar to the RP literature, the four SP studies defined severity of disease in different ways: (1) baseline health-related quality of life (HR-QoL) [47]; (2) impact on survival relative to quality of life [48]; and (3) expected survival and HR-QoL [46, 49]. Efficacy and cost-effectiveness also differed depending on which variables were included and which were held constant. For example, one study [47] modelled ICER only, but not what the ratio consisted of, while another [48] modelled ICER, incremental QALYs per patient and number of patients separately, but not incremental cost per patient. The other studies only modelled disaggregated components of the ICER, holding either total cost [49] or total patients [46] constant.

3.3.3 Outcome Variable, Econometric Methods and Model Specification

Three studies (three analyses) modelled hypothetical decisions using a binary logit model [47–49]; two studies (two analyses) used an MNL model [46, 48]; and one study (one analysis) used a mixed logit (MXL) model [46], where two studies estimated more than one class of model. The choice of model was influenced by the number of alternatives and type of hypothetical questions asked. For example, committee members were asked to either recommend or reject one technology at a time [47], make a forced choice between two alternative technologies [49], make a flexible choice between two technologies or neither [46], or both a forced choice between two technologies and then a flexible choice between both, one or none [48]. By construction, the selection of variables is specified a priori in SP studies.

3.4 Factors Influencing Decisions

Table 1 presents the direction of effect that key content factors had on making a positive recommendation compared with ‘not recommended’ across studies by committee (see the electronic supplementary material for comparison with restricted decisions). Few studies defined ‘technology

characteristics’ (such as innovation, pharmaceutical, preventative), which were mostly not important. Across the studies, the RP literature suggested that the main factors driving decisions were clinical efficacy (in terms of statistical significance in the trial evidence or the acceptance of clinical effect/evidence by the committee) and, where applicable, cost-effectiveness (i.e. ICER). That is, committees were more likely to recommend a health technology when there was good evidence that it is effective, but less likely when health gains were costlier (per unit of health). The SP literature generally found clinical efficacy (defined as size and certainty of effect), cost-effectiveness and equity factors were all important. However, due to differences in the types of factors considered by different committees, understanding what drives specific committees is most relevant. Therefore, this section of the review will focus on the PBAC, NICE and the AWMSG, which were the most commonly studied and for which a comparison between SP and RP studies was possible.

3.4.1 PBAC (Australia)

Overall the SP and RP literature consistently found the PBAC were less likely to make a positive recommendation as budget impact, ICER and ICER uncertainty increased, but more likely when the committee accepted or trusted the clinical evidence and for more severe diseases. The first two findings were supported by all five RP studies, and the latter three were supported by the subset of RP studies that modelled them (see Table 1). These findings were also supported indirectly by the SP evidence, which found committee members were less likely to recommend a pharmaceutical with lower incremental survival or quality of life, higher incremental cost or higher uncertainty [46]. All else constant, this corresponds to preferences for smaller ICERs, lower budget impact and confidence or acceptance of the evidence.

The RP studies also consistently indicated that the PBAC was more likely to recommend a pharmaceutical it had previously considered (all else constant), which may be evidence of a bargaining process [23, 25, 26]. In contrast, the effect of there being no alternative therapy available was unclear given two studies found no effect [23, 26] while two found a counter-intuitive negative effect [25, 27]. The negative effect observed in these two studies may have been driven by the inclusion of decisions based on CMA, given such decisions always have an alternative therapy (to cost-minimise) and also had a high probability of being recommended. Therefore, a strong negative correlation between no alternative therapy and positive recommendation in CMAs may have dominated the overall effect.

Table 1 Impact of key content factors on positive recommendations (reference category of not recommended) by committee, study and analysis

Committee	Study/analysis	Therapeutic value										
		Higher level of clinical evidence	Direct vs indirect clinical evidence	Greater relevance of clinical evidence	Larger quantity of clinical evidence	Greater quality of clinical evidence	Positive or size of clinical efficacy	Accept vs reject clinical evidence				
PBAC	Harris et al. [23]	A1 ○	○	○	○	○	○	○	○	○	+	
	Chim et al. [24]	A2 ×	×	○	×	×	×	×	×	×	+	
		A1 ○		×	×	×	○				×	
	Mauskopf et al. [25]	A2 ○		×	×	×	○					
		A3 ×		×	×	×	×					
		+		○							+	
	Harris et al. [26]										+	
	Karikiotis et al. [27]	A1	×	×	×	×	×	×	×	×	+	
	Whitty et al. [46]	A2	×	×	×	×	×	×	×	+	+	
		A1									+	
A2									+	+		
NICE	Devlin and Parkin [28]											
	Dakin et al. [29]	A1	○				+/○	○	○	○		
	Mauskopf et al. [30]	A2	○					+/○	○	○		
		A1 ○	○			×					+	
	Cerri et al. [31]	A2 ○	○		×						+	
		×	×					×				
	Dakin et al. [32]	A1					○					
	A2						○					
	AWMSG	Tappenden et al. [47]										
		Linley and Hughes [33]	A1			-		○				
A2								○				
Linley and Hughes [48]		A1									+	

Table 1 continued

Committee	Study/analysis	Therapeutic value							
		Higher level of clinical evidence	Direct vs indirect clinical evidence	Greater relevance of clinical evidence	Larger quantity of clinical evidence	Greater quality of clinical evidence	Positive or size of clinical efficacy	Accept vs reject clinical evidence	
CADTH-CEDAC	Rocchi et al. [35]		○	○			+	+	
CADTH-pCODR	Skedgel [49]						+	+	
CT	Le Pen et al. [38]	A1					+	+	
HIRA	Park et al. [40]	A2	○		×		+	+	
HIRA-NHIS	Kim et al. [41]	A1						○	
NCPE	Schmitz et al. [42]	A1				+		×	
CVZ	Cerri et al. [39]	A2	+		×	+	+		
AOTM	Niewada et al. [34]	A3				+		+	
SMC	Charokopou et al. [43]		×		×		×		
TLV	Svensson et al. [44]								
CRM	Pauwels et al. [37]		×		°		+	×	
CMS	Chambers et al. [36]				○		+		
NICE/SMC/CVZ/CT	Cerri et al. [45]	A1	+		○		+		
NICE/SMC	Cerri et al. [45]	A2	+		+/○		○		

Table 1 continued

Committee	Study/analysis	Economic value					Disease characteristics					
		CUA and/ or CEA vs other	CMA or similar vs other	Larger ICER or INMB	Larger ICER or uncertainty	Accept vs reject economic evidence	Larger budget or relative cost	Cancer vs other disease area	Larger prevalence of disease	Greater severity of disease	No alternative vs alternative therapy	Orphan vs non-orphan disease
PBAC	Harris et al. [23]	A1	-	-	-	-/0	-	-	+	○		
		A2	-	-	-	○	-	-	+	×		
Chim et al. [24]		A1	-	○	-						×	
	Mauskopf et al. [25]	A2	-	-	-							
		A3	×	-	- ^a					- ^a		
Harris et al. [26]		A1	×	×	-	-	-		+	○		
	Karikios et al. [27]	A2	×	×	- ^a					- ^a		
Whitty et al. [46]		A1			-				+			
		A2			-				+			
NICE	Devlin and Parkin [28]	A1	○		-				+	+		
	Dakin et al. [29]	A2	-		○					○		
Mauskopf et al. [30]		A1	-		-					×		
		A2	×		-					×		
Cerri et al. [31]		A1			-					×		×
	Dakin et al. [32]	A2			-					×		×
Tappenden et al. [47]		A1			-					○		○
		A2			-					+		○
AWMSG	Linley and Hughes [33]	A1			○					○		○
		A2			×					×		×
Linley and Hughes [48]		A1			-							
		A2			-							
CADTH-CEDAC	Rocchi et al. [35]		×	-	×	○						○
	Skedgel [49]											
CADTH-pCODR		A1			-							
		A2			-							
CT	Le Pen et al. [38]	A1										
		A2										
HIRA	Park et al. [40]											

Table 1 continued

Committee	Study/analysis	Economic value				Disease characteristics						
		CUA and/or CEA vs other	CMA or similar vs other	Larger ICER or INMB	Larger ICER uncertainty	Accept vs reject economic evidence	Larger budget or relative cost	Cancer vs other disease area	Larger prevalence of disease	Greater severity of disease	No alternative vs alternative therapy	Orphan vs non-orphan disease
HIRA-NHIS	Kim et al. [41]	A1 ○	+								○	○
NCPE	Schmitz et al. [42]	A2 ×	+								×	×
		A1		-	×		×		×		×	×
		A2		-		-	×		×		×	×
CVZ	Ceri et al. [39]	A3		-	×		×		×		×	×
			×	×			○			×	×	×
AOTM	Niewada et al. [34]					+	×					
SMC	Charokopou et al. [43]		+				×			×		×
TLV	Svensson et al. [44]			-					+			
CRM	Pauwels et al. [37]	×		×			×					×
CMS	Chambers et al. [36]											
NICE/SMC/CVZ/CT	Ceri et al. [45]	A1					×				○	+
NICE/SMC	Ceri et al. [45]	A2 ×		-			×			○	×	×

A1, A2 and A3 refer to the unique analyses presented in the studies which met the inclusion criteria

Symbols illustrate the relationship between the variable and a positive recommendation (vs not recommended): + a positive relationship ($p \leq 0.05$); - a negative relationship ($p \leq 0.05$); ○ not important ($p > 0.05$); × defined by study, but not included in multivariate model. Symbols may represent one or more variable reflecting the same or similar interpretations associated with the content factors included in the table

AOTM Agencia Ocenii Medycznych (Agency for Health Technology Assessment), AWMG All Wales Medicines Strategy Group, CADTH Canadian Agency for Drugs and Technologies in Health, CART classification and regression tree, CEA cost-effectiveness analysis, CEDAC Canadian Expert Drug Advisory Committee, CMA cost-minimisation analysis, CMS Centres for Medicare and Medicaid Services, CRM Commission de Remboursement des Médicaments (The Commission of Reimbursement of Medicines), CT Commission de la Transparence (Transparency Commission), CUA cost-utility analysis, CVZ College Voor Zorgverzekering (Health Care Insurance Board), HIRA Health Insurance Review and Assessment service, ICER incremental cost-effectiveness ratio, INMB incremental net monetary benefit, NCPE National Centre for Pharmacoeconomics, NHIS National Health Insurance Service, NICE National Institute for Health and Care Excellence, PBAC Pharmaceutical Benefits Advisory Committee, pCODR pan-Canadian Oncology Drug Review, SMC Scottish Medicines Consortium, TLV Tandvårds-och läkemedelsförmånsverket (Dental and Pharmaceutical Benefits Agency)

^a Conditional effect within CART

^b Modelled as clinical uncertainty

^c Observational studies only

3.4.2 NICE (England & Wales)

ICER was the only factor consistently identified across the SP and RP literature which influenced NICE decisions, where a higher ICER was associated with a lower chance of recommendation (see Table 1). Evidence across other factors was mixed, with a number of differences between SP and RP findings. For example, SP evidence indicated NICE committee members were also less likely to make positive recommendations as ICER uncertainty increased, but more likely where no alternative treatment existed or for more severe diseases [47]. The first two findings were consistent with one RP study [28], but others found no effect for ICER uncertainty [32] or no alternative therapy [29, 32]. The only RP study to model severity of disease found no effect [32]. However, the 'end of life' criteria adopted by NICE in 2009 may be associated with severity. Although one study [30] found the end of life criteria had no effect based mostly on decisions before it was applicable, another study [32] found it had a positive effect after 2009.

There was also mixed evidence across the RP literature for other factors. For example, cancer was either prioritised [32] or deprioritised [30], budget impact either had a negative effect [30] or no effect [29], and the number of appraisals considered simultaneously was either important [32] or not [31]. Involvement by patient advocates had no effect in the studies where it was modelled [29, 32]. Overall, the dominance of the ICER as the main factor influencing decisions in the RP studies may be due to explicit ICER thresholds used by NICE [52].

3.4.3 AWMSG (Wales)

No important factors were consistently identified across the two studies of the AWMSG. The SP study by Linley and Hughes [48] indicated committee members were less likely to recommend a pharmaceutical as ICER or ICER uncertainty increased and more likely as clinical efficacy (in QALYs) increased or where no alternative therapy existed. While the latter two factors were not modelled in the RP study also by Linley and Hughes [33], sample characteristics may explain the lack of effect for cost-effectiveness (defined as INMB) and uncertainty (presence of probabilistic sensitivity analysis). For example, the sample ($n = 60$) included a large proportion of decisions for orphan drugs ($n = 20$) and HIV ($n = 12$), for which cost-effectiveness may be less sensitive, as well as decisions based on CMA ($n = 8$), for which ICER and probabilistic sensitivity analysis are not considered.

3.5 Impact of Important Factors on Decisions

Figure 3 presents the marginal effects of ICER [increase of A\$10,000 (Australian dollars) or £4800], budget impact

(increase of A\$10 million or £12.34 million per annum) and whether the committee raised concerns or had issues with the clinical evidence across studies of the PBAC and NICE from a reference point of 50%. This can be interpreted as the change in the probability of receiving a positive recommendation from 50% for the given change in each factor. For example, an effect size of -10% corresponds to a reduction in the probability that the committee would recommend a technology from 50% to 40%. The figure illustrates that a comparable change in ICER had a similar impact on decisions across studies of the PBAC and NICE, with two exceptions. The effect estimated from Devlin and Parkin [28] was larger than that from other RP studies, but this may be due to the small sample ($n = 33$) and inclusion of both cost per QALY and cost per LY in the ICER variable. The impact was also relatively larger based on SP data in Tappenden et al. [47]. In contrast to ICER, the marginal effects of budget impact differed across studies and committees. While not important for NICE, the role of budget impact may have changed over time for the PBAC, with greater importance in more recent decisions. For example, Harris et al. [23] found no effect between 1994 and 2004, whereas others found a large effect based on decisions after 2005 [24–27]. These findings are consistent with the official positions of these committees, given budget impact is not currently considered by NICE for decision making [52], but became a relevant factor for the PBAC after 2006 [53, 54]. When either committee had an issue or raised concern with the clinical evidence, the probability of making a positive recommendation tended towards zero. The size of this effect was approximately equivalent to reducing the ICER by A\$38,000/QALY for the PBAC or £15,000/QALY for NICE (excluding the two outliers).

4 Discussion

This review has highlighted a growing and diverse literature on how preferences of HTA decision makers for factors considered in funding decisions have been investigated and the results of such studies. While the majority of the literature has focused on analysing past decisions, there were both RP and SP studies for the PBAC, NICE and AWMSG. Despite differences in variable definitions, the SP literature generally found clinical efficacy (size and uncertainty of effect), cost-effectiveness and equity factors (such as disease severity) were important for these committees, as well as budget impact for the PBAC. While these findings were supported by the RP studies of the PBAC, the RP literature generally found funding decisions for other committees were driven by clinical efficacy (acceptance of clinical evidence) and, where applicable, cost-effectiveness. Whether equity factors were

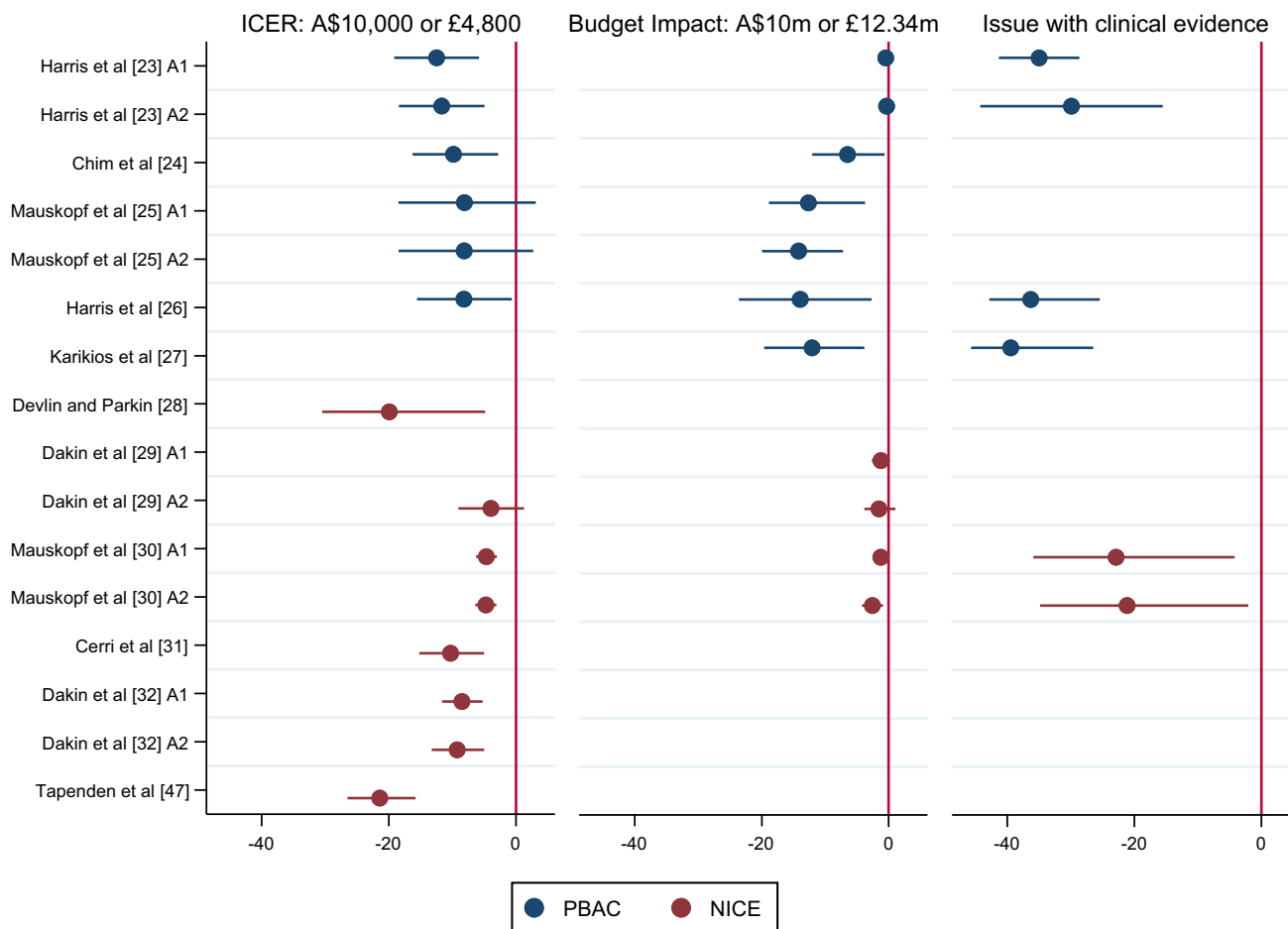


Fig. 3 Marginal effect on the probability of positive recommendation for the PBAC and NICE, from a reference point of 50%. Average purchasing power parity (2004–2016) = 0.48 British pounds (£) to 1 Australian dollar (A\$). Average relative size of health budget in the United Kingdom to Australia (2004–2016) = 2.57. When not reported, the mean values of categorical variables from PBAC decisions were based on data provided by Harris et al. [23] and Harris

et al. [26]. The estimated effects for ICER and/or budget impact from Chim et al. [24], Mauskopf et al. [25] and Harris et al. [26] reflect the average effect across multiple categories. A\$ Australian dollars, *ICER* incremental cost-effectiveness ratio, *m* million, *NICE* National Institute for Health and Care Excellence, *PBAC* Pharmaceutical Benefits Advisory Committee

important for past funding decisions was unclear, particularly for NICE. This review also found that the relative impact of key factors investigated (ICER, budget impact and acceptance of clinical evidence) were generally consistent across RP studies of the PBAC and NICE, despite expected changes over time. Both committees appeared to place considerable importance on having trust in the clinical data relative to the other factors. However, these findings must be interpreted with due consideration to the many methodological differences and data constraints across the literature, as well as differences in the way individual HTA agencies and committees are organised and operate.

HTA decisions in the ‘real world’ are based on the principles of substantive and procedural justice in the context of information asymmetry, incomplete information and uncertain future events [55]. Other issues faced by committees include rationality with past decisions,

difficulties combining individual opinions and strategic bargaining via repeated decisions [26, 55]. Controlling for these and other contextual factors in quantitative analysis is challenging and has rarely been attempted in the RP literature. In contrast, stated or hypothetical decisions are made in the absence of such real-world considerations, and may more closely reflect decision makers’ actual priorities, albeit for a restricted set of criteria. Therefore, while SP and RP studies aim to answer the same research questions, there may be a fundamental difference in what SP and RP data captures. This is important and may partly explain the differences observed across the two literatures, as well as the difficulties in validating SP data with RP data [48].

The interpretation of many of the RP studies as reflecting preferences for the trade-off between improvements in health, increased cost and other factors is also problematic for a number of reasons. First, as the majority

of RP studies relied on exploratory statistical selection techniques to inform which factors were modelled, criteria stated to be assessed and important to individual committees were often dropped from the analysis. Consequently, such models often included a combination of variables without clear rationale, such as ICER uncertainty without ICER [33]. Second, most RP studies modelled decisions from CMAs (i.e. decisions without ICERs) and CEAs (i.e. decisions with ICERs) together, often to preserve sample size. This is an issue because there is no clear trade-off between different factors in decisions based on CMA as the benefits and costs of the new and comparator technologies should be equivalent. That is, provided the committee accepts that new technology offers the same health benefit and the cost is the same (or cheaper), there would be no obvious reason to reject. However, researchers have defined and assigned various equity factors (including severity of disease, cancer, or no alternative therapy) to technologies appraised using CMA, leading to unpredictable effects in those studies.

In addition, the RP literature has largely not accounted for the sequential nature of HTA decisions given clinical evidence is considered before cost-effectiveness and other factors [6, 7]. One of the two studies included in the review which estimated a hierarchical model (i.e. CART analysis) found evidence that cost-effectiveness was only important after the PBAC first accepted the clinical evidence [27]. This is also consistent with two-way analyses in other included studies [24, 30, 34, 35, 40] indicating that acceptance of the clinical evidence is *almost* necessary, but not sufficient for a positive recommendation. Presumably, other processes such as price negotiations or risk-sharing agreements are used to support decisions where the clinical evidence is poor. However, hierarchical CART analysis assumes that every factor is considered in a step-wise fashion, which may not be realistic, given some committees appear to trade-off between health gains, equity factors and cost simultaneously at some point [4]. Attempts to account for sequential decision making in standard parametric choice models are also limited by the number of interaction terms that can be estimated from small samples. Therefore, a combination of the two approaches may be preferred, for example, using sample selection models or by placing a greater focus on key subgroups (i.e. excluding CMAs and decisions where the clinical evidence was accepted).

Data quality was another key limitation for many RP studies. Given the commercially sensitive nature of the relevant information, researchers have had to interpret any documents made available by committees in order to define variables (which may not have actually been considered). A further challenge is that relevant factors are not defined by most HTA committees and these could be subject to change over time. Small sample sizes may too introduce

other potential problems such as low statistical power and insufficient variation within variables.

Of course there are also many differences in the objectives of individual committees as well as their procedural guidelines and stated decision criteria, which helps explain the diversity of factors explored [1, 5]. For example, not all HTA committees consider 'value for money' using CEA. Funding decisions by some committees may be based on clinical criteria alone and price controls, with or without consideration of other factors. The range of stated factors which should influence decisions also varies greatly in different countries, from human value, solidarity and need in Sweden to public health, own responsibility and rarity in the Netherlands [56–59]. It is important to note that most of the committees identified in this review do evaluate cost-effectiveness along with other types of factors and provide (at least some) public justification for their decisions. There are other HTA committees for which analyses of decisions have not been undertaken in the literature and to which the findings of this review may not apply. However, these results are expected to prove useful for key stakeholders (including patient groups, industry and committees themselves) within the countries considered in this review, as well as to policy makers in jurisdictions considering establishing or reviewing HTA procedures [60, 61].

It is worth mentioning that this review does not intend to summarise all the factors that may potentially affect HTA funding decisions. Rather, this review has focused on the quantitative literature and in particular on identifying important factors and quantifying their impact on HTA decisions after controlling for the influence of other factors. The qualitative literature on this topic can provide further context to these important decisions [4–7].

Future research using RP data should take greater account of (1) the real-world setting; (2) the estimation of confirmatory rather than exploratory analyses; (3) the exclusion of decisions without a clear trade-off between health gains, equity and cost (such as decisions based on CMA); and (4) the sequential and then simultaneous nature of HTA decision making. The fast-changing nature of this field must also be considered. While a push for greater transparency of committee decisions may improve data quality over time, the rise in confidential risk-sharing arrangements may have the opposite effect as the actual prices and ICERs are hidden [62]. Similarly, the role of uncertainty may change with the rise of coverage with evidence decisions and after-market reviews [63, 64]. In contrast, future SP studies could investigate more complex relationships by holding fewer variables constant, especially given committee members are highly informed participants. Preferences for the less common or emerging special cases (including 'rule of rescue', 'end of life', children or rare diseases) which are difficult to study using

past decisions should also be explored using SP data. Finally, as the number of both past decisions and committee members increase over time, larger samples in both RP and SP studies may allow the specification of interaction effects. In particular, interactions between ICER and equity factors could be used to better understand differences in ICER premiums or thresholds for health technologies [19].

5 Conclusions

By comparing the methods and findings of published studies that have investigated preferences of HTA committees for factors considered in past and hypothetical funding decisions, this review found that trust in the clinical evidence and, where applicable, cost-effectiveness are key factors for decision makers. The relative importance of having that confidence in the evidence was approximately equivalent to reducing the ICER by A\$38,000/QALY for the PBAC or £15,000/QALY for NICE. Although the role of equity and other factors varied across individual committees as expected, their role also varied across studies of the same committee. The diversity in some of these findings was likely due to the many methodological differences across the literature, including differences in how factors were defined and analysed. Given significant interest from stakeholders in understanding what influences decision outcomes and considerable resources expended in this area, work on this topic is expected to continue to expand. We hope that this review and suggestions for further research to address current gaps in the literature will help in this ongoing endeavour to understand what drives HTA decision making.

Acknowledgements The author team would like to thank Dennis Petrie for his helpful suggestions and valuable advice in drafting the review.

Author Contributions All authors designed the systematic review. PG and YG applied the selection criteria to the identified studies; EL adjudicated any differences. PG and SZ extracted and synthesised the data. PG drafted the manuscript, with input from the other authors. PG acts as guarantor for the paper and accepts full responsibility for the conduct of the review and decision to publish.

Compliance with Ethical Standards

Funding The research was supported by a National Health and Medical Research Council (NHMRC) Project Grant (APP1047788): “Societal and decision maker preferences for priority setting in health care resource allocation”.

Conflict of interest EL declares no conflict of interest. PG, YG and SZ declare they are contracted through their respective universities to evaluate submissions for listing of pharmaceuticals on the Pharmaceutical Benefits Scheme in Australia. Neither they nor their spouses, partners or children have any other financial or non-financial

relationships that may be relevant to the submitted work. The content of this paper does not reflect the views of the Australian Government Department of Health, the Pharmaceutical Benefits Advisory Committee or its sub-committees.

Data Availability Statement Data sharing is not applicable as no datasets were generated or analysed. All findings were based on the published studies included in this review.

References

1. Drummond M. Twenty years of using economic evaluations for drug reimbursement decisions: what has been achieved? *J Health Polit Policy Law*. 2013;38(6):1081–102.
2. Gu Y, Lancsar E, Ghijben P, Butler JR, Donaldson C. Attributes and weights in health care priority setting: a systematic review of what counts and to what extent. *Soc Sci Med*. 2015;146:41–52.
3. Dolan P, Shaw R, Tsuchiya A, Williams A. QALY maximisation and people’s preferences: a methodological review of the literature. *Health Econ*. 2005;14(2):197–208.
4. Vuorenkoski L, Toiviainen H, Hemminki E. Decision-making in priority setting for medicines—a review of empirical studies. *Health Policy*. 2008;86(1):1–9.
5. Stafinski T, Menon D, Philippon DJ, McCabe C. Health technology funding decision-making processes around the world. *Pharmacoeconomics*. 2011;29(6):475–95.
6. Erntoft S. Pharmaceutical priority setting and the use of health economic evaluations: a systematic literature review. *Value Health*. 2011;14(4):587–99.
7. Niessen LW, Bridges J, Lau BD, Wilson RF, Sharma R, Walker DG, et al. Assessing the impact of economic evidence on policymakers in health care—a systematic review. 2012.
8. Fischer KE. A systematic review of coverage decision-making on health technologies—evidence from the real world. *Health Policy*. 2012;107(2):218–30.
9. Moher D, Liberati A, Tetzlaff J, Altman DG, The PG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097.
10. Clark S, Weale A. Social values in health priority setting: a conceptual framework. *J Health Org Manag*. 2012;26(3):293–316.
11. Rotter JS, Foerster D, Bridges JF. The changing role of economic evaluation in valuing medical technologies. *Expert Rev Pharmacoecon Outcomes Res*. 2012;12(6):711–23.
12. Guindo LA, Wagner M, Baltussen R, Rindress D, van Til J, Kind P, et al. From efficacy to equity: literature review of decision criteria for resource allocation and healthcare decisionmaking. *Cost Effect Resour Allocat*. 2012;10(1):9.
13. Kaufman RL. Comparing effects in dichotomous logistic regression: a variety of standardized coefficients. *Soc Sci Quart* 1996;90–109.
14. OECD Data. 2016; Available from: <https://data.oecd.org/>.
15. Higgins JPT, Green S. *Cochrane handbook for systematic reviews of interventions*: Wiley; 2011.
16. Anis AH, Guh D, Wang X-H. A dog’s breakfast: prescription drug coverage varies widely across Canada. *Med Care*. 2001;39(4):315–26.
17. Segal L, Dalziel K, Mortimer D. Fixing the game: are between-silo differences in funding arrangements handicapping some interventions and giving others a head-start? *Health Econ*. 2010;19(4):449–65.
18. Chambers JD, Morris S, Neumann PJ, Buxton MJ. Factors predicting medicare national coverage: an empirical analysis. *Med Care*. 2012;50(3):249–56.

19. Schilling C, Mortimer D, Dalziel K. Using CART to Identify thresholds and hierarchies in the determinants of funding decisions. *Med Decis Making*. 2016;0272989X16638846.
20. Al MJ, Feenstra T, Brouwer WB. Decision makers' views on health care objectives and budget constraints: results from a pilot study. *Health Policy*. 2004;70(1):33–48.
21. Koopmanschap MA, Stolk EA, Koolman X. Dear policy maker: have you made up your mind? A discrete choice experiment among policy makers and other health professionals. *Int J Technol Assess Health Care*. 2010;26(02):198–204.
22. Lee H-J, Bae E-Y. Eliciting preferences for medical devices in South Korea: a discrete choice experiment. *Health Policy*. 2017;121(3):243–9.
23. Harris AH, Hill SR, Chin G, Li JJ, Walkom E. The role of value for money in public insurance coverage decisions for drugs in Australia: a retrospective analysis 1994–2004. *Med Decis Making*. 2008.
24. Chim L, Kelly PJ, Salkeld G, Stockler MR. Are cancer drugs less likely to be recommended for listing by the Pharmaceutical Benefits Advisory Committee in Australia? *Pharmacoeconomics*. 2010;28(6):463–75.
25. Mauskopf J, Chirila C, Masaquel C, Boye KS, Bowman L, Birt J, et al. Relationship between financial impact and coverage of drugs in Australia. *Int J Technol Assess Health Care*. 2013;29(01):92–100.
26. Harris A, Li JJ, Yong K. What can we expect from value-based funding of medicines? A retrospective study. *Pharmacoeconomics*. 2016;34(4):393–402.
27. Karikios DJ, Chim L, Martin A, Nagrial A, Howard K, Salkeld G, et al. Is it all about price? Why requests for government subsidy of anticancer drugs were rejected in Australia. *Int Med J*. 2017;47(4):400–7.
28. 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.
29. Dakin HA, Devlin NJ, Odeyemi IA. “Yes”, “No” or “Yes, but”? Multinomial modelling of NICE decision-making. *Health Policy*. 2006;77(3):352–67.
30. Mauskopf J, Chirila C, Birt J, Boye KS, Bowman L. Drug reimbursement recommendations by the National Institute for Health and Clinical Excellence: have they impacted the National Health Service budget? *Health Policy*. 2013;110(1):49–59.
31. Cerri KH, Knapp M, Fernandez J-L. Decision making by NICE: examining the influences of evidence, process and context. *Health Econ Policy Law*. 2014;9(02):119–41.
32. Dakin H, Devlin N, Feng Y, Rice N, O'Neill P, Parkin D. The influence of cost-effectiveness and other factors on nice decisions. *Health Econ*. 2015;24(10):1256–71.
33. Linley WG, Hughes DA. Reimbursement decisions of the All Wales Medicines Strategy Group. *Pharmacoeconomics*. 2012;30(9):779–94.
34. Niewada M, Polkowska M, Jakubczyk M, Golicki D. What influences recommendations issued by the agency for health technology assessment in Poland? A glimpse into decision makers' preferences. *Value Health Region Issue*. 2013;2(2):267–72.
35. Rocchi A, Miller E, Hopkins RB, Goeree R. Common drug review recommendations. *Pharmacoeconomics*. 2012;30(3):229–46.
36. Chambers JD, Chenoweth M, Cangelosi MJ, Pyo J, Cohen JT, Neumann PJ. Medicare is scrutinizing evidence more tightly for national coverage determinations. *Health Aff*. 2015;34(2):253–60.
37. Pauwels K, Huys I, De Nys K, Casteels M, Simoens S. Predictors for reimbursement of oncology drugs in Belgium between 2002 and 2013. *Expert Rev Pharmacoecon Outcomes Res*. 2015;15(5):859–68.
38. Le Pen C, Priol G, Lilliu H. What criteria for pharmaceuticals reimbursement? *Eur J Health Econ*. 2003;4(1):30–6.
39. Cerri KH, Knapp M, Fernandez J-L. Public funding of pharmaceuticals in the Netherlands: investigating the effect of evidence, process and context on CVZ decision-making. *Eur J Health Econ*. 2014;15(7):681–95.
40. Park SE, Lim SH, Choi HW, Lee SM, Kim DW, Yim EY, et al. Evaluation on the first 2 years of the positive list system in South Korea. *Health Policy*. 2012;104(1):32–9.
41. Kim E-S, Kim J-A, Lee E-K. National reimbursement listing determinants of new cancer drugs: a retrospective analysis of 58 cancer treatment appraisals in 2007–2016 in South Korea. *Expert Rev Pharmacoecon Outcomes Res*. 2017 (just-accepted).
42. Schmitz S, McCullagh L, Adams R, Barry M, Walsh C. Identifying and revealing the importance of decision-making criteria for health technology assessment: a retrospective analysis of reimbursement recommendations in Ireland. *Pharmacoeconomics*. 2016;34(9):925–37.
43. Charokopou M, Majer IM, de Raad J, Broekhuizen S, Postma M, Heeg B. Which factors enhance positive drug reimbursement recommendation in Scotland? A retrospective analysis 2006–2013. *Value Health*. 2015;18(2):284–91.
44. Svensson M, Nilsson FO, Arnberg K. Reimbursement decisions for pharmaceuticals in Sweden: the impact of disease severity and cost effectiveness. *Pharmacoeconomics*. 2015;33(11):1229–36.
45. Cerri KH, Knapp M, Fernandez J-L. Untangling the complexity of funding recommendations: a comparative analysis of health technology assessment outcomes in four European countries. *Pharm Med*. 2015;29(6):341–59.
46. Whitty JA, Scuffham PA, Rundle-Thiele SR. Public and decision maker stated preferences for pharmaceutical subsidy decisions. *Appl Health Econ Health Policy*. 2011;9(2):73–9.
47. Tappenden P, Brazier J, Ratcliffe J, Chilcott J. A stated preference binary choice experiment to explore NICE decision making. *Pharmacoeconomics*. 2007;25(8):685–93.
48. Linley WG, Hughes DA. Decision-makers' preferences for approving new medicines in Wales: a discrete-choice experiment with assessment of external validity. *Pharmacoeconomics*. 2013;31(4):345–55.
49. Skedgel C. The prioritization preferences of pan-Canadian Oncology Drug Review members and the Canadian public: a stated-preferences comparison. *Curr Oncol*. 2016;23(5):322.
50. ISPOR Global Health Care Systems Road Map. 2016; Available from: <http://www.ispor.org/htaroadmaps/france.asp>.
51. Bach F. Model-consistent sparse estimation through the bootstrap. arXiv preprint [arXiv:09013202](https://arxiv.org/abs/09013202). 2009.
52. Social value judgements: principles for the development of NICE guidance, 2nd edn: National Institute for Health and Clinical Excellence; 2008.
53. Guidelines for the pharmaceutical industry on preparations of submissions to the Pharmaceutical Benefits Advisory Committee. Australian Government Department of Health and Ageing; 2002.
54. Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.0). Australian Government Department of Health and Ageing; 2006.
55. Fischer KE, Leidl R. Analysing coverage decision-making: opening Pandora's box? *Eur J Health Econ*. 2014;15(9):899–906.
56. Franken M, Stolk E, Scharringhausen T, de Boer A, Koopmanschap M. A comparative study of the role of disease severity in drug reimbursement decision making in four European countries. *Health Policy*. 2015;119(2):195–202.
57. Gulácsi L, Rotar AM, Niewada M, Löblová O, Rencz F, Petrova G, et al. Health technology assessment in Poland, the Czech

- Republic, Hungary, Romania and Bulgaria. *Eur J Health Econ.* 2014;15(1):13–25.
58. Stafinski T, Menon D, Davis C, McCabe C. Role of centralized review processes for making reimbursement decisions on new health technologies in Europe. *Clin Econ Outcomes Res CEOR.* 2011;3:117.
59. Franken M, le Polain M, Cleemput I, Koopmanschap M. Similarities and differences between five European drug reimbursement systems. *Int J Technol Assess Health Care.* 2012;28(4):349–57.
60. ISPOR Initiative on US Value Assessment Frameworks. 2017; Available from: <https://www.ispor.org/ValueAssessmentFrameworks/Index>.
61. Kaló Z, Gheorghe A, Huic M, Csanádi M, Kristensen FB. HTA implementation roadmap in Central and Eastern European countries. *Health Econ.* 2016;25(S1):179–92.
62. Robertson J, Walkom EJ, Henry DA. Transparency in pricing arrangements for medicines listed on the Australian Pharmaceutical Benefits Scheme. *Aust Health Rev.* 2009;33(2):192–9.
63. Lexchin J. Coverage with evidence development for pharmaceuticals: a policy in evolution? *Int J Health Serv.* 2011;41(2):337–54.
64. Parkinson B, Sermet C, Clement F, Crausaz S, Godman B, Garner S, et al. Disinvestment and value-based purchasing strategies for pharmaceuticals: an international review. *Pharmacoeconomics.* 2015;33(9):905–24.