

Policy

Cite this article: Efthymiadou O, Kanavos P (2021). Determinants of Managed Entry Agreements in the context of Health Technology Assessment: a comparative analysis of oncology therapies in four countries. *International Journal of Technology Assessment in Health Care* 37, e31, 1–7. <https://doi.org/10.1017/S0266462321000039>

Received: 3 July 2020

Revised: 21 December 2020

Accepted: 7 January 2021

Key words:

Risk sharing; Managed Entry Agreements; Patient access schemes; HTA decision making; HTA determinants

Author for correspondence:

Olina Efthymiadou, E-mail: a.efthymiadou@lse.ac.uk

Determinants of Managed Entry Agreements in the context of Health Technology Assessment: a comparative analysis of oncology therapies in four countries

Olina Efthymiadou  and Panos Kanavos

Medical Technology Research Group, Department of Health Policy, London School of Economics, Houghton Street, London WC2A 2AE, UK

Abstract

Background. Managed Entry Agreements (MEAs) are increasingly used to address uncertainties arising in the Health Technology Assessment (HTA) process due to immature evidence of new, high-cost medicines on their real-world performance and cost-effectiveness. The literature remains inconclusive on the HTA decision-making factors that influence the utilization of MEAs. We aimed to assess if the uptake of MEAs differs between countries and if so, to understand which HTA decision-making criteria play a role in determining such differences.

Methods. All oncology medicines approved since 2009 in Australia, England, Scotland, and Sweden were studied. Four categories of variables were collected from publicly available HTA reports of the above drugs: (i) Social Value Judgments (SVJs), (ii) Clinical/Economic evidence submitted, (iii) Interpretation of this evidence, and (iv) Funding decision. Conditional/restricted decisions were coded as Listed With Conditions (LWC) other than an MEA or LWC including an MEA (LWCMEA). Cohen's κ -scores measured the inter-rater agreement of countries on their LWCMEA outcomes and Pearson's chi-squared tests explored the association between HTA variables and LWCMEA outcomes.

Results. A total of 74 drug-indication pairs were found resulting in $n = 296$ observations; 8 percent ($n = 23$) were LWC and 55 percent ($n = 163$) were LWCMEA. A poor-to-moderate agreement existed between countries ($-.29 < \kappa < .33$) on LWCMEA decisions. Cross-country differences within the LWCMEA sample were partly driven by economic uncertainties and largely driven by SVJs considered across agencies.

Conclusions. A set of HTA-related variables driving the uptake of MEAs across countries was identified. These findings can be useful in future research aimed at informing country-specific, "best-practice" guidelines for successful MEA implementation.

Background and Objectives

Over the past decade, the continuous market entry of new therapies, which are either high volume—for treating many patients (i.e., antidiabetic agents) or high cost—for a single treatment course (i.e., oncology therapies), has escalated pharmaceutical spending (1). It was recently reported that pharmaceutical spending accounts for almost 20 percent of the total health expenditure in OECD countries, and because funding from governments and social insurance schemes plays the largest role in pharmaceutical purchasing, this rise bears significant implications for the budget of health systems (2).

The growing healthcare expenditure poses pressures for pharmaceutical manufacturers to demonstrate real-world value for money beyond that of safety and efficacy and simultaneously for national healthcare payers to engage in strategic pricing and reimbursement policies that ensure patients' access to new therapies while optimizing budget impact (3;4). Although most new products are assessed as part of Health Technology Assessment (HTA) processes in many countries, the data available on the cost-effectiveness of high-cost therapies, particularly in oncology, are severely lacking at the time of product launch (5). Uncertainties arise due to the often immature evidence available from controlled trials on the real-world clinical outcomes of newly launched pharmaceuticals, meaning that the benefits of a new product cannot be fully estimated at drug launch; uncertainties may be present around treatment eligibility of patient subgroups, generalizability of trial results to clinical practice, the use of surrogate outcome measures instead of "hard" end points and subsequent transferability of surrogate outcomes used in trials to real-world studies (6). As these challenges can lead to delayed reimbursement decisions and patient access, manufacturers and payers are seeking ways to collaboratively manage the market entry of new pharmaceutical products and mitigate risks related to premature evidence (7;8); one way to achieve this has been through the introduction of contractual arrangements

© The Author(s), 2021. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

between the two parties, referred to as Managed Entry Agreements (MEAs) or Risk Sharing Agreements (RSAs).

MEAs are used in many countries worldwide and primarily in Europe, in accordance with country-specific governance and preferences around evidence requirements and evaluation for new medicines (9). Cross-country differences in HTA assessment requirements have led to a well-documented disparity in the respective risk-sharing practices followed by countries (6). A review of MEAs in the EU showcased that only for two drug-indication pairs in the whole sample, an MEA was applied to all six study countries, and even between these, there was variation in the type of MEAs applied (10).

Despite the growing interest of healthcare systems and manufacturers in the use of MEAs over the past decade, there is still a knowledge gap in the drivers of this variation, partially due to a lack of transparency in the negotiation from both parties (11).

Even though the literature has concluded that countries indeed differ in their MEA implementation practices, with MEAs being highly specific to the HTA decision-making context in which they operate, the current body of relevant literature arises mainly from secondary evidence and remains largely descriptive in nature (8;12–14). Therefore, we aimed to (i) make a methodological contribution to existing research on determining whether the uptake of MEAs differs between countries and (ii) if so, to understand whether specific HTA decision-making uncertainties and considerations play some role in determining such differences.

Methods

Sample Selection

A retrospective analysis of HTA appraisals on all oncology drugs (i.e., all L01 molecules, based on the Anatomical Therapeutic Chemical classification) which obtained regulatory approval by the European Medicines Agency (EMA) and by the Therapeutic Goods Administration (TGA) in Australia between 1 January 2009 and 15 June 2018 (at the drug-indication pair level) in Australia (AUS), England (ENG), Scotland (SCOT), and Sweden (SE). Oncology was selected as our study therapeutic class because it has been documented to be the therapeutic class with the largest proportion (38 percent) of implemented MEAs; it is also the therapeutic class where MEAs continue to be increasingly implemented (6). Study countries were selected because they all implement MEAs, they all have long-established HTA policies and processes to guide their coverage decisions, they have both a publicly available list of MEAs and publicly available HTA reports that provide a sufficient level of information for the purposes of this analysis, and they use similar criteria in their decision-making processes (i.e., clinical and/or cost-effectiveness), allowing for comparability across agencies (15).

The sample used for this analysis comprises a small part of a larger sample of drugs studied for a different, broader project on HTA. Nevertheless, the aim/scope of that project is not related to that of our study and neither is the methodology we used for data analysis and management. The only common aspects between the two studies relate to the overarching framework used for data collection, as well as the classification and validation of dimensions studied, as described below.

Methodological Framework for Data Collection

The methodology underpinning the data collection process was based on the literature (15;16) where it is suggested that the

final outcome of an HTA appraisal is driven by (i) the clinical and cost-effectiveness evidence submitted (i.e., clinical trial design and end points, safety, economic models, and comparators) and (ii) the interpretation of this evidence, influenced by the perception of uncertainty around this evidence, by setting-specific preferences on risk and other social value considerations.

For the purposes of our analysis, a simplified methodological framework was adopted based on the assumption that the impact of the clinical and economic evidence submitted (Stage 1) on the final decision outcome (Stage 4) is captured through the respective uncertainties that this evidence has raised or not raised (Stage 2). Therefore, the final HTA outcome (Stage 4) is a function of the uncertainties raised (Stage 2) and other social value and system-specific considerations (Stage 3) (Figure 1).

The HTA process was divided in four different stages corresponding to: (1) *the evidence submitted and appraised* (e.g., trial type, clinically meaningful end points; response rate/disease progression/safety end points, comparators, Incremental Cost-Effectiveness Ratio (ICER) range and economic model used), (2) *the interpretation of this evidence/uncertainties raised* (i.e., clinical and economic evidence-related uncertainties around clinical benefit and study/research design and those around economic modeling and cost-effectiveness, respectively), (3) *Social Value Judgments (SVJs) and HTA system-specific considerations* (i.e., additional dimensions of societal value that a new technology adds, beyond its clinical evidence/benefit and cost-effectiveness such as the innovation and administration advantage it offers, value dimensions specific to the disease area the technology addresses, such as its severity, rarity, and unmet need, or whether it is a condition toward the end of life, where the benefit of a treatment is valued more highly, and/or system-specific considerations such as the use of a single or multiple technology appraisal (MTA) in England that shape decision-making processes for each study country), and (4) *the final decision outcome*, classified as (i) L = List (i.e., positive HTA recommendation), (ii) LWC = List with conditions, where the technology has been accepted with restrictions but which are not classified as MEAs (e.g., a product should be used in a subpopulation of its licensed indication, and/or it should be used in a second line or higher line of therapy, and/or it should be used in a specific dose only, and/or it requires monitoring, and/or it requires prescription by a specialist), (iii) LWCMEA = Listed with conditions/restrictions including, among others (if any), at least one classified as an MEA (e.g., simple discount, free stock, rebate, patient access scheme, commercial access agreement coverage with evidence development and/or additional data collection), and (iv) DNL = Do not list (i.e., negative HTA recommendation). We preferred a four-category outcome variable over the three-outcome variable traditionally used in the HTA literature (i.e., listed, listed with conditions, rejected) as it better reflects the multiple coverage options available when studying conditional/restricted HTA decisions. As L and DNL decisions would, by definition, not lead to some kind of a condition/restriction, for the purposes of this analysis, we studied only those drug-indication pairs with a conditional/restricted recommendation decision (i.e., LWC or LWCMEA).

Data Collection

SVJs and uncertainties were classified and defined based on the literature (Supplementary Appendix, Table 1) (12;15;17;18), and the classification was also discussed and validated between the authors and external referees. Data on the above stages per

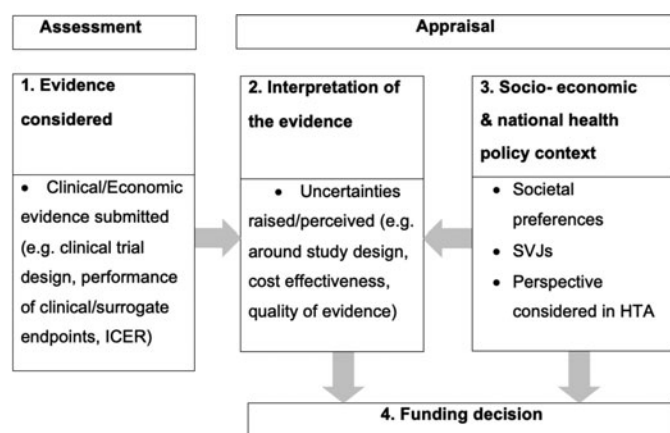


Figure 1. Methodological framework on the analysis of the HTA process and variables included therein. Note: ICER, Incremental Cost-Effectiveness Ratio; SVJs, Social Value Judgments. Source: The authors based on the literature (15;16).

drug-indication pair in all study countries were extracted only from the official, publicly available HTA appraisals, which were published in the Web sites of the respective HTA bodies (i.e., PBAC (AUS), NICE (ENG), SMC (SCOT)), TLV (SE)) (Supplementary Appendix, Table 2); other relevant sources of data, such as the county councils' group on new drug therapies (NLT) in Sweden were not searched. Where needed, searches were conducted in local languages (English and Swedish) to improve the accuracy and comprehensiveness of the extraction. The data extracted were fed into a database stratified by an HTA agency to describe and classify MEAs across the respective HTA bodies and ultimately facilitate data analyses. Data collection was undertaken between June and November 2018 and only the final HTA recommendation reports available for the drug-indication pairs studied were used for data collection (Supplementary Appendix, Table 3).

Data Analysis

For the first part of our hypothesis (i.e., determining whether the uptake of MEAs differs between countries), Pearson's chi-squared (χ^2) test of independence was used to test for differences in the restricted outcome (LWC or LWCMEA) between agencies. Cohen's kappa scores (κ) of cross-country agreement levels were also measured as an additional robustness check.

Agreement between agencies was measured based on whether conditional/restricted HTA outcomes across agencies included a form of MEA as a condition (or not); this allowed a comparison of the observed inter-rater agreement with the agreement expected by chance, ranging from poor ($\kappa = 0$) to perfect ($\kappa = 1$), with negative values of κ corresponding to cases where the inter-rater agreement was even less than that expected by chance (19).

Finally, for the second part of our research question (i.e., understand whether specific HTA decision-making uncertainties and considerations play some role in driving divergent LWCMEA decisions between countries), variables under the HTA appraisal (Figure 1) were initially analyzed by means of descriptive statistics, including percentages and crosstabulations (Excel® 2013). Assuming that all categories of uncertainties and SVJs were applicable to all drug-indication pairs studied, these were treated and coded as binary variables based on whether they have been raised and considered (or not), respectively, in the HTA report for each drug-indication pair. As such, bivariate

analyses were also performed, using Pearson's chi-squared (χ^2) test of independence, to identify which of these variables drive differences between the LWC and LWCMEA sample overall, and across agencies. For the former comparison, where uncertainty and SVJ dimensions had small sample sizes (i.e., five or less observations), the Fisher's exact test was also performed as a robustness check. The SPSS® (v.24.0) was used to perform statistical tests and measure inter-rater agreement.

Results

Sample Characteristics

A total of 74 molecules were studied across the four countries, corresponding to a total sample of $n = 296$ drug-indication pairs. Of these, 7 percent ($n = 21$) were Listed (L), 9.5 percent ($n = 28$) were Not Listed (DNL), and 63 percent ($n = 186$) were Listed with at least one type of condition/restriction (LWC or LWCMEA). Other outcomes included drug-indication pairs that were "Not Assessed" (10 percent, $n = 29$), "Not submitted" (4 percent, $n = 12$), or formed only an "Advice or Health Economic Report" (6.5 percent, $n = 20$) (Table 1).

Of the conditional/restricted HTA outcomes ($n = 186$), 88 percent ($n = 163$) were LWCMEA overall, but only 17 percent ($n = 32$) were LWCMEA across all countries for the same molecules. This was further emphasized by the Cohen's kappa scores measuring the level of inter-rater agreement between agencies across their LWCMEA outcomes. The scores ranged from -0.29 to 0.33 (Table 2), demonstrating a poor-to-moderate agreement. More precisely, it was shown that only SCOT and SE had a moderate agreement, whereas the rest of the countries had a poor or even negative agreement between them. Cross-country differences in MEA utilization were further strengthened by results of the χ^2 test, which showed statistically significant differences between the study countries in terms of their LWCMEA recommendation decisions ($P < .05$).

Clinical and Economic Uncertainties

Among all clinical uncertainties raised, those that seemed to differ distinctly in proportion between LWC and LWCMEA were population generalizability (13% vs. 31%, respectively), followed by relevance to clinical practice (22% vs. 37%) and study design (35% vs. 51%) (Table 1). Uncertainties around all other clinical evidence aspects were raised at a nearly equal proportion between the LWC and LWCMEA sample, for example, clinical evidence submitted (48% vs. 47%, respectively), clinical benefit (65% vs. 72%), and clinical comparator (35% vs. 29%). Overall, it was shown by χ^2 tests, and where applicable, by the Fisher's exact test, that there were no statistically significant differences between LWC and LWCMEA groups in terms of clinical uncertainties (Table 1).

Clinical uncertainties did not drive any statistically significant differences between countries within the LWC sample. Nevertheless, in the LWCMEA sample, agencies differed significantly in raising clinical uncertainties around *study design* ($P < .05$), *clinical comparator* ($P < .05$), *population generalizability* ($P < .001$), and *relevance to clinical practice* ($P < .0001$) (Figure 2).

Differences between the LWC and LWCMEA groups were more prominent when studying economic uncertainties. Those that underpinned significant differences included *utilities* (4% vs. 42%, $P < .0001$; Fisher's exact significance), followed by

Table 1. Differences in HTA variables studied between conditional/restricted recommendation decisions (LWC and LWCMEA) and their respective *P* values

All recommendation decisions (<i>n</i> = 296)		
List (L)	21 (7%)	
List with conditions		
LWC	23 (8%)	
LWCMEA	163 (55%)	
Do not list (DNL)	28 (9.5%)	
Not assessed	29 (10%)	
Not submitted	12 (4%)	
Economic report	20 (6.5%)	
Conditional/Restricted recommendation decisions (<i>n</i> = 186)		
Number of assessments per country (* <i>P</i> < .05)		
	LWC	LWCMEA
England (NICE)	1 (4%)	58 (36%)
Australia (PBAC)	12 (52%)	38 (23%)
Scotland (SMC)	3 (13%)	51 (31%)
Sweden (TLV)	7 (31%)	16 (10%)
Social Value Judgments		
	LWC	LWCMEA
Rarity (<i>P</i> = .07)		
Considered	7 (30%)	25 (15%)
Not considered	16 (70%)	138 (85%)
Disease severity (<i>P</i> = .33)		
Considered	7 (30%)	67 (42%)
Not considered	16 (70%)	96 (59%)
Unmet need (<i>P</i> = .33)		
Considered	13 (56.5%)	109 (67%)
Not considered	10 (43.5%)	54 (33%)
Administration advantage (* <i>P</i> < .01)		
Considered	2 (9%)	54 (33%)
Not considered	21 (91%)	109 (67%)
Innovation (* <i>P</i> < .001)		
Considered	2 (9%)	86 (53%)
Not considered	21 (91%)	77 (47%)
Special Considerations (i.e., end-of-life criteria) (* <i>P</i> < .05)		
Considered	7 (30%)	92 (56%)
Not considered	16 (70%)	71 (43%)
Clinical uncertainties		
	LWC	LWCMEA
Clinical benefit (<i>P</i> = .47)		
Raised	15 (65%)	118 (72%)
Not raised	8 (35%)	45 (28%)
Study design (<i>P</i> = .14)		
Raised	8 (35%)	83 (51%)
Not raised	15 (65%)	80 (49%)

Relevance to clinical practice (<i>P</i> = .14)		
Raised	5 (22%)	61 (37%)
Not raised	18 (78%)	102 (63%)
Population generalizability (<i>P</i> = .08)		
Raised	3 (13%)	50 (31%)
Not raised	20 (87%)	113 (69%)
Clinical comparator (<i>P</i> = .60)		
Raised	8 (35%)	48 (29%)
Not raised	15 (65%)	115 (71%)
Clinical evidence (<i>P</i> = .95)		
Raised	11 (48%)	77 (47%)
Not raised	12 (52%)	86 (53%)
Economic uncertainties		
	LWC	LWCMEA
Cost-effectiveness (* <i>P</i> < .01)		
Raised	9 (39%)	117 (72%)
Not raised	14 (61%)	46 (28%)
Utilities (* <i>P</i> < .01)		
Raised	1 (4%)	68 (42%)
Not raised	22 (96%)	95 (58%)
Costs (<i>P</i> = .31)		
Raised	8 (35%)	75 (46%)
Not raised	15 (65%)	88 (54%)
Modeling (<i>P</i> = .95)		
Raised	13 (57%)	118 (72%)
Not raised	10 (43%)	45 (28%)
Model type (<i>P</i> = .90)		
Raised	1 (4%)	8 (5%)
Not raised	22 (96%)	155 (95%)
Economic comparator (* <i>P</i> < .05)		
Raised	0 (0%)	35 (22%)
Not raised	23 (100%)	128 (78%)

Note: LWC, Listed With Criteria; LWCMEA, Listed With Criteria, including an MEA; NICE, National Institute of Clinical Excellence; SMC, Scottish Medicines Consortium; PBAC, Pharmaceutical Benefits Advisory Committee; TLV, The Dental and Pharmaceutical Benefits Agency.

economic comparator (0% vs. 22%, *P* < .01; Fisher's exact significance) and cost-effectiveness (39% vs. 72%, *P* < .01) (Table 1).

Furthermore, uncertainties around utilities (*P* < .0001), economic comparator (*P* < .01), cost-effectiveness (*P* < .01), modeling (*P* < .0001), model type (*P* < .05), and costs included (*P* < .0001) also drove differences between agencies within the LWCMEA group (Figure 2). Finally, only cost-effectiveness (*P* < .05) generated statistically significant differences among the agencies within the LWC sample.

Social Value Judgments

There was no statistical difference in the likelihood that most categories of SVJs were considered for drugs that were listed on the

Table 2. *K* scores (*k*, [95 percent CI]) of inter-rater agreement in the commonly assessed and conditional/restricted HTA outcomes across countries

	England (NICE)	Scotland (SMC)	Australia (PBAC)	Sweden ^a (TLV)
England (NICE)	-	-.03 [-.07;.01]	-.05 [-.14;.04]	-.12 [-.32;.08]
Scotland (SMC)	-	-	-.05 [-.14;.04]	.33 [-.17;.83]
Australia (PBAC)	-	-	-	-.29 [-.51;-.07]
Sweden (TLV)	-	-	-	-

Note: NICE, National Institute of Clinical Excellence; SMC, Scottish Medicines Consortium; PBAC, Pharmaceutical Benefits Advisory Committee; TLV, The Dental and Pharmaceutical Benefits Agency.

^aScores were generated only with the mutual drug-indication pairs that were both commonly assessed and listed with conditions (LWC or LWCMEA) among all study countries (i.e., 15 molecules commonly assessed between England, Scotland, Australia, and Sweden and 34 molecules commonly assessed and listed with conditions between England, Scotland, and Australia).

basis of LWC versus LWCMEA. Exceptions to this were *innovation* (9% vs. 53%, $P < .0001$; Fisher's exact significance), *administration advantage* (9% vs. 33%, $P < .05$; Fisher's exact significance), and *special considerations* (30% vs. 56%, $P < .05$) (Table 1). In contrast, within the LWCMEA sample, statistically significant differences were observed across countries in the likelihood of considering most SVJs, including *unmet need* ($P < .0001$), *special considerations* ($P < .0001$), *impact on society* ($P < .01$), *impact on Quality of Life (QoL)* ($P < .0001$), *impact on emotional and functional burden* ($P < .0001$, respectively), *severity of disease* ($P < .0001$), *innovation* ($P < .0001$), and *administration advantage* ($P < .0001$) (Figure 2). Finally, only the last three SVJs seemed to drive statistically significant ($P < .05$) cross-country differences in the LWC sample too.

Discussion

We demonstrated significant disparities in the conditional/restricted recommendations across all cancer drugs appraised by four HTA agencies between 2009 and 2018. More precisely, we demonstrated a poor level of agreement in MEA implementation across countries as indicated by the kappa scores. Our results suggest that the countries followed different strategies in dealing with the risk/uncertainty arising from the respective evidence submitted by manufacturers on new oncology therapies.

Diverging MEA outcomes between countries were influenced heavily by economic evidence uncertainties including those around cost-effectiveness, utilities, and costs included in the economic model, highlighting agency-specific preferences on cost-effectiveness thresholds and evidentiary requirements for economic modeling. Similar findings around the importance of economic considerations, and notably, the criterion of cost-effectiveness in determining the final HTA recommendation, have been described elsewhere (19–21).

Clinical evidence uncertainties were less influential than economic toward listing a drug with an MEA; this was not surprising because in many cases it has been demonstrated that uncertainties around the strength of clinical evidence and benefit often lead to rejections commonly across agencies, without allowing any

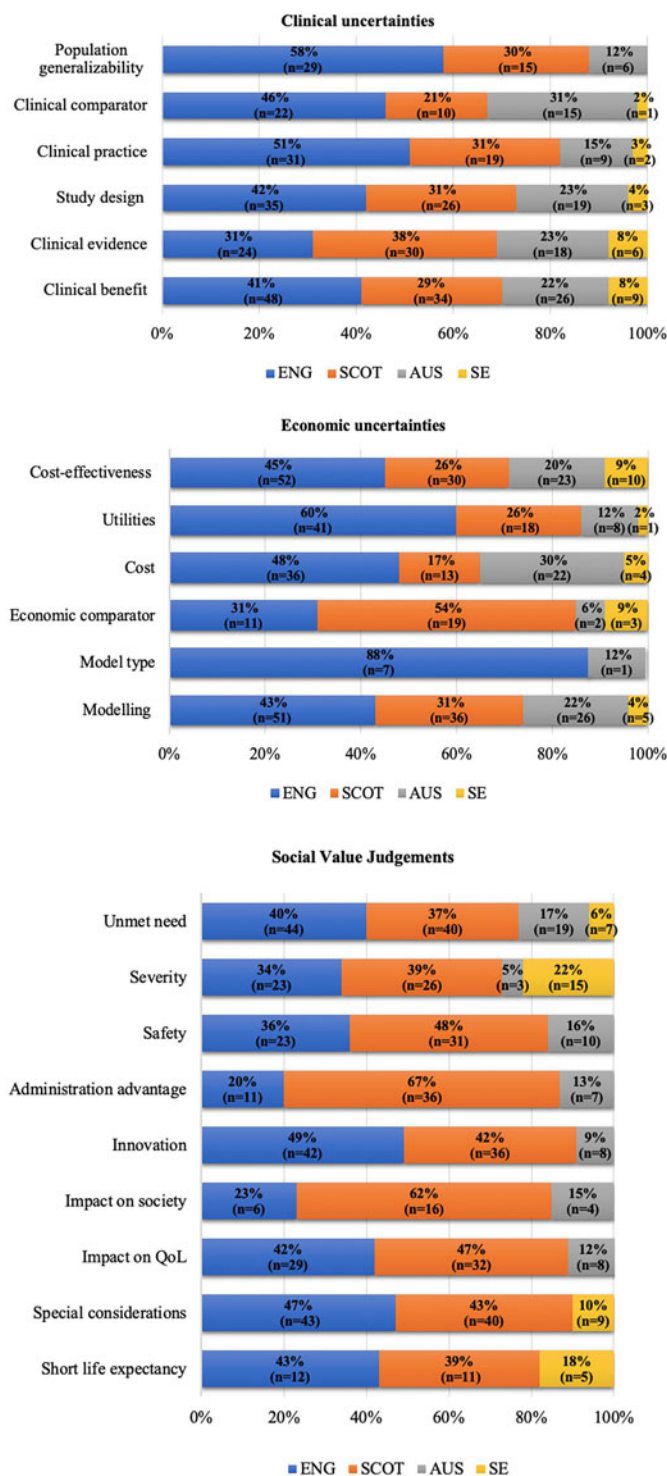


Figure 2. Cross-country variation in clinical/economic uncertainties and SVJs raised and considered respectively by HTA agencies for drug-indication pairs approved with MEAs. Note: ENG, England; SCOT, Scotland; AUS, Australia; SE, Sweden; QoL, Quality of Life. Source: The authors.

flexibility for conditions or funding negotiations (15;19). However, clinical uncertainties related to setting-specific characteristics (i.e., relevance of the technology in question and of the clinical comparator to the country/region-specific clinical practice, and/or the generalizability of trial population to the setting-specific population) were found to play a role in cross-country

variation within the LWCMEA sample. This finding confirms that some agencies might place a greater emphasis on evidence related to clinical practice and trial population compared with other agencies (15). It follows that uncertainties around these factors may also play a role in the uptake of risk-sharing negotiations across specific countries.

In terms of SVJs, our findings suggest that most social value considerations determined cross-country differences within the LWCMEA group only, with the exception of innovation, administration advantage, and severity of disease, which underpinned variation within the LWC group too. This observation highlights that considerations around innovation and burden of disease might be crucial in allowing greater flexibility toward funding with conditions/restrictions in general, whereas considerations around impact on QoL, societal impact, emotional/functional burden, as well as other special considerations (i.e., end-of-life criteria) could be influential specifically toward funding with an MEA as a restriction.

Indeed, the HTA literature has recognized that factors such as the burden of disease the treatment addresses, aspects of the treatment's innovation level, but also the wider socioeconomic implications of the treatment largely affect the perceived value of new medical technologies (22). Similarly, a number of setting-specific decision modulators (e.g., the SMC modifiers, the NICE's end-of-life criteria, and the human dignity principle for TLV) can contribute to a greater flexibility toward the acceptance of uncertainty or higher and uncertain ICERs (18). Nevertheless, as shown in Figure 2, the extent to which the above factors are considered across countries in their LWCMEA recommendation decisions can fluctuate significantly, whereas even in countries where these factors are taken into account, they are not necessarily reported in their assessment processes.

Ultimately, this links to discussions about the role of MEAs in the implementation of Value-Based Pricing (VBP) policies through negotiations that enable weight-adjustment of cost-effectiveness thresholds for new medicines that tackle diseases with a higher burden, demonstrate greater therapeutic innovation, and/or have wider societal benefits, such that they reflect any of these additional elements. For example, experience from TLV has shown that in Sweden, risk-sharing agreements indeed complement the VBP system for out-patient drugs and enable stakeholders to mitigate different types of uncertainty (23). Of course, greater clarity on the long-term outcomes of MEAs is also key to understanding the feasibility of MEA negotiations as a tool for the efficient enactment of VBP policies.

To the best of our knowledge, this is the first systematic analysis that confirms cross-country differences in the uptake of MEAs and provides an understanding of the HTA decision-making variables that might influence such differences. Similar, largely descriptive studies have also identified differences in the design and implementation of MEAs across countries but still lack an in-depth analysis and transparency around the HTA determinants of MEAs (6;8;11;13;24). To date, there have been no best-practice guidelines in MEA implementation, and only a few scientific papers suggest some related principles, such as the proposal by KCE in Belgium for good practice in (performance-based) MEAs (25). As such, results from this analysis contribute to shedding light on the rationale/strategies behind the implementation of MEAs across countries and, therefore, facilitate policy relevant research on the creation of implementation guidelines and/or regulations on "risk-sharing" policies. Finally, providing a transparent, evidence-based description of the HTA decision-

making variables that can typically influence an approval with MEAs could be applied in practice by policy makers to facilitate/accelerate HTA decision making and, therefore, allow for more timely reimbursement decisions and consequently more timely access to new, high-cost medicines.

Limitations

This research contributes to an improved understanding of the potential factors that drive conditional/restricted HTA recommendations with versus without an MEA and why these two outcomes might differ significantly between countries for the same drug, presenting with similar clinical evidence across countries. Nevertheless, our findings should be interpreted with caution, because there are certain limitations in our analysis that have hindered the accuracy and robustness of our results.

Firstly, variables under Stage 1 of the HTA process (Figure 1), such as the type of clinical evidence submitted or the ICER submitted, have not been included in this analysis based on the assumption that their influence on determining an LWC or LWCMEA outcome will be captured through their respective uncertainties and whether these were raised or not in the decision-making process. Because variables under both stages have been found to have an impact on the final HTA recommendation (15;23), future analyses could include these variables as a robustness check.

Additionally, as budget impact is taken into account only by the Pharmaceutical Benefits Advisory Committee (PBAC), a budget impact-specific uncertainty was not considered due to the lack of comparability across countries. Nevertheless, any budget impact-related concerns raised by the PBAC would have already been reflected in the "cost"- and "modeling"-related uncertainties.

Furthermore, as the calculation of the κ scores required the assumption of paired observations to be met, we performed the calculation only on those drug-indication pairs that were assessed by all four agencies, thus reducing the available sample size significantly. As these analyses could have been more robust if the sample size was increased, it is suggested that a future replication of this study augments the sample size through the inclusion of molecules from additional therapeutic areas. Similarly, as our analysis covered medicines reimbursed only at the national level, future analyses could also account for technologies negotiated at the hospital level.

Finally, advanced statistical modeling was not used at this stage, as our aim was to understand which of the variables play at least some role in driving key differences in the conditional/restricted HTA recommendation decisions between countries. However, in subsequent analyses, the key variables identified herein can be included in a multinomial logistic regression model to also understand their level of impact on divergencies between the conditional/restricted HTA recommendation decisions and the types of MEAs in place across countries.

Overall, it is important that the findings reported here are interpreted with caution, given that the data collected and analyzed were sourced from publicly accessible reports that may, in some instances, represent amended versions of the assessment process to preserve manufacturers' commercial sensitivities and as such, may not present an absolute reflection of the committee's considerations.

Conclusions and the Way Forward

MEAs are implemented globally and particularly in oncology, to address uncertainties arising from the high cost and simultaneous

immature clinical evidence of new, innovative pharmaceuticals. We showed that the uptake of MEAs across countries for the same drugs might differ substantially, and it is subject to setting-specific evidentiary requirements on economic modeling, the comparators, costs, and utilities included therein but primarily also subject to preferences on social value considerations, such as the socioeconomic and QoL impact of the treatment appraised, as well as setting the specific burden of disease. A better understanding of the criteria that determine MEA utilization across countries is fundamental for future research aimed at informing country-specific, “best-practice” guidelines for successful MEA negotiations.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/S0266462321000039>.

Acknowledgments. We are grateful to the following experts for their invaluable collaboration in the data collection and validation process: Anna-Maria Fontrier, Mackenzie Mills, and Erica Visintin.

Funding. The authors received no specific funding for this work.

References

1. Permanand G, Bak Pedersen H. Managing new premium-priced medicines in Europe. *J Pharm Policy Pract.* 2015;**8** (Suppl 1):K2.
2. OECD. *Health at a glance 2019: OECD indicators.* Paris: OECD Publishing. 2019 [cited 2020 Sep 10]. Available from: <https://www.oecd-ilibrary.org/docserver/4dd50c09-en.pdf?expires=1601653288&id=id&accname=guest&checksum=F2FAFFEF26F158A09DCDA2E104FBEF8F> doi:10.1787/4dd50c09-en
3. McCabe CJ, Bergmann L, Bosanquet N, Ellis M, Enzmann H, von Euler M, et al. Market and patient access to new oncology products in Europe: A current, multidisciplinary perspective. *Ann Oncol.* 2009;**20**:403–12.
4. Taylor RS, Drummond MF, Salkeld G, Sullivan SD. Inclusion of cost effectiveness in licensing requirements of new drugs: The fourth hurdle. *BMJ.* 2004;**329**:972–5.
5. Garrison Jr LP, Towse A, Briggs A, De Pouvourville G, Grueger J, Mohr PE, et al. Performance-based risk-sharing arrangements—Good practices for design, implementation, and evaluation: Report of the ISPOR good practices for performance-based risk-sharing arrangements task force. *Value Health.* 2013;**16**:703–19.
6. Ferrario A, Kanavos P. *Managed entry agreements for pharmaceuticals: The European experience.* Brussels, Belgium: EMI-net. 2013 [cited 2018 Oct 15]. Available from: http://eprints.lse.ac.uk/50513/1/_Libfile_repository_Content_Ferrario%2C%20A_Ferrario_Managed_%20entry_%20agreements_2013_Ferrario_Managed_%20entry_%20agreements_2013.pdf
7. Wilsdon T, Barron A. Charles river associates. Managed Entry Agreements in the context of medicines adaptive pathways to patients. Final Report. 2016 [cited 2018 Jan 28]. Available from: <http://adaptmart.eu/wp-content/uploads/2016/12/CRA-MEA-in-the-context-of-MAPPs-Final-Report-16-December-2016-STC.pdf>
8. Carlson JJ, Sullivan SD, Garrison LP, Neumann PJ, Veenstra DL. Linking payment to health outcomes: A taxonomy and examination of performance-based reimbursement schemes between healthcare payers and manufacturers. *Health Policy.* 2010;**96**:179–90.
9. Pritchett L, Wiesinger A, Faria-Billinton E, Brown A, Murray G, Stoor L, et al. A review of guidelines and approaches to performance-based risk-sharing agreements across the UK, Italy and the Netherlands. *Value Health.* 2015;**18**:A568.
10. Pauwels K, Huys I, Vogler S, Casteels M, Simoens S. Managed Entry Agreements for oncology drugs: Lessons from the European experience to inform the future. *Front Pharmacol.* 2017;**8**:171.
11. Piatkiewicz TJ, Traulsen JM, Holm-Larsen T. Risk-sharing agreements in the EU: A systematic review of major trends. *Pharmacoecon Open.* 2017;**2**:109–23.
12. Ferrario A, Kanavos P. Dealing with uncertainty and high prices of new medicines: A comparative analysis of the use of managed entry agreements in Belgium, England, the Netherlands and Sweden. *Soc Sci Med.* 2015;**24**:39–47.
13. Morel T, Arickx F, Befrits G, Siviero P, van der Meijden C, Xoxi E, et al. Reconciling uncertainty of costs and outcomes with the need for access to orphan medicinal products: A comparative study of managed entry agreements across seven European countries. *Orphanet J Rare Dis.* 2013;**8**:198.
14. Adamski J, Godman B, Ofierska-Sujkowska G, Osińska B, Herholz H, Wendykowska K, et al. Risk sharing arrangements for pharmaceuticals: Potential considerations and recommendations for European payers. *BMC Health Serv Res.* 2010;**10**:153.
15. Nicod E, Kanavos P. Developing an evidence-based methodological framework to systematically compare HTA coverage decisions: A mixed methods study. *Health Policy.* 2016;**120**:35–45.
16. Cerri KH, Knapp M, Fernández JL. Decision making by NICE: Examining the influences of evidence, process and context. *Health Econ Policy Law.* 2014;**9**:119–41.
17. Angelis A, Kanavos P. Multiple criteria decision analysis (MCDA) for evaluating new medicines in health technology assessment and beyond: The advance value framework. *Soc Sci Med.* 2017;**188**:137–56.
18. Nicod E, Kanavos P. Scientific and social value judgements for orphan drugs in health technology assessment. *Int J Technol Assess Health Care.* 2016;**32**:218–32. Epub 2016/09/14.
19. Nicod E. Why do health technology assessment coverage recommendations for the same drugs differ across settings? Applying a mixed methods framework to systematically compare orphan drug decisions in four European countries. *Eur J Health Econ.* 2016;**18**:715–30.
20. Maynou L, Cairns J. What is driving HTA decision-making? Evidence from cancer drug reimbursement decisions from 6 European countries. *Health Policy.* 2019;**123**:130–39. Epub 2018/11/20. doi:10.1016/j.healthpol.2018.11.003.
21. Maynou-Pujolras L, Cairns J. Why do some countries approve a cancer drug and others don't? *J Cancer Policy.* 2015:21–5. [Health Policy. 2019;123(2):130–9. Epub 2018/11/20.
22. Kanavos P, Angelis A. Multiple criteria decision analysis for value based assessment of new medical technologies: A conceptual framework. 2013 [cited 2020 Apr 3]. Available from: http://www.lse.ac.uk/LSEHealthAndSocialCare/pdf/LSEHealthworkingpaperseries/LSEH-WP33_final.pdf
23. Andersson E, Svensson J, Persson U, Lindgren P. Risk sharing in managed entry agreements - A review of the Swedish experience. *Health Policy.* 2020;**124**:404–10. Epub 2020/02/18.
24. Gonçalves FR, Santos S, Silva C, Sousa G. Risk-sharing agreements, present and future. *Ecancermedalscience.* 2018;**12**:823.
25. Vogler S, Paris V, Panteli D. Policy brief 30: Ensuring access to medicines: How to redesign pricing, reimbursement and procurement? World Health Organization 2018 (acting as the host organization for, and secretariat of, the European Observatory on Health Systems and Policies). 2018 [cited 2020 Apr 14]. Available from: http://www.euro.who.int/_data/assets/pdf_file/0009/379710/PolicyBrief_AUSTRIA_PB30_web_13082018.pdf