HEALTH ECONOMICS,
POLICY and LAW

ARTICLE

Why drugs fail health technology assessment: a comparative analysis of health technology assessment rejections across seven OECD countries

Filippos Papadopoulos[1], Erica Visintin[1,2], Ilias Kyriopoulos[1] and Panos Kanavos[1,2]

[1]Department of Health Policy, London School of Economics and Political Science, London, UK and [2]LSE Health – Medical Technology Research Group, London School of Economics and Political Science, London, UK
Corresponding author: Panos Kanavos; Email: p.g.kanavos@lse.ac.uk

Abstract
While a substantial amount of evidence exists on factors associated with positive health technology assessment (HTA) outcomes, the evidence on the same regarding rejections is scarce. Using a proprietary dataset of HTA outcomes in seven Organisation for Economic Co-operation and Development (OECD) countries, we empirically examine the factors associated with HTA rejections and study the magnitude of inter-agency differences in technology appraisals. Data were extracted from HTA reports between 2009 and 2020. The primary outcome was the probability of rejection, which was examined with respect to several regulatory, disease-related, evidence (clinical and economic) and unaddressed uncertainty variables. Multivariate logistic regression analysis was used. Out of N = 1,405 HTA assessments, the rejection rate was 12.9% (n = 181). Significant predictors of HTA rejection were submissions for drugs with cancer or orphan indications (but not both), low quality of evidence and the presence of uncertainties surrounding clinical benefit, cost-effectiveness, and economic model utility inputs. Systematic differences between agencies in their propensity for rejecting the same drugs were revealed, particularly in relation to cancer and rare diseases. Despite the low rejection rate, our findings suggest that it is critical to improve quality of evidence, focus on risk mitigation strategies as a means of reducing the impact of uncertainties and share HTA practices across borders to increase consistency in decision-making.

1. Introduction

While regulatory approval is a necessary condition for the placement of new therapies on the market, it is no longer sufficient in ensuring timely patient access to healthcare innovations; health technology assessment (HTA) is a critical step that new therapies must undergo before a healthcare system makes them available to patients at public expense. The role of HTA is critical in ensuring that valuable innovations, offering significant additional clinical benefit to patients, are included in the benefits health systems provide.

Over time, many HTA agencies have been criticised in that they lack transparency, clear evaluation criteria, and flexible approaches for the assessment of drugs with atypical evidence submissions (e.g. orphan or ultra-orphan drugs) (Angelis, Lange, and Kanavos 2018). Despite calls to streamline and standardise HTA principles and methods across agencies or countries (O'Mahony 2019; Kristensen 2012), there remain significant differences in how agencies are

structured (Martelli *et al.* 2007; Allen *et al.* 2013) and in the methods they employ to reach decisions (Mathes *et al.* 2013; Vreman *et al.* 2020; Akehurst *et al.* 2017). This heterogeneity in HTA principles, processes and methods has led to disparities in coverage decisions between agencies often within the borders of the same country (Nicod and Kanavos 2012; Fischer, Heisser, and Stargardt 2016; Salas-Vega, Bertling, and Mossialos 2016). This is understandable and, in some way, expected, since healthcare systems vary, the access to services they provide varies, and priorities differ. Nevertheless, understanding whether commonalities exist among clinical, economic and contextual factors influencing coverage decisions can help approximate decision processes and promote the perceived rationality of HTA processes amongst citizens.

Evidence suggests that a significant proportion of new drug submissions to HTA agencies receive a positive recommendation (Kanavos, Visintin and Gentilini, 2023), but a fair number are rejected. In principle, rejections of new drugs by HTA agencies safeguard the public from spending scarce resources on novel drugs with little or no added clinical value, but considering that the same drug can be approved in one setting but be rejected in another, raises broader equity concerns.

There are multiple dimensions of 'value' considered either formally or informally across national HTA bodies. These dimensions can be roughly grouped into considerations of the drug's added therapeutic benefit (focus of agencies that do not use a cost-effectiveness criterion like Germany's IQWiG and France's HAS), comparative clinical and economic value (focus of agencies requiring demonstration of cost-effectiveness like England's NICE), budgetary impact and perceived societal importance (Beletsi *et al.* 2018). When reviewing the clinical and economic evidence on a new drug, decision-makers make value judgements, guided (at least in principle) by their agency's evidence evaluation framework. The relative importance placed on these different elements of value differ considerably across agencies.

The divergence in HTA recommendations across countries is well-documented in the literature. Most comparative empirical studies to date have employed descriptive statistical methods (Fischer, Heisser, and Stargardt 2016; Grepstad and Kanavos 2015; Allen *et al.* 2017; Vreman *et al.* 2020), with some attempting to further explain differences in HTA outcomes through case studies (Nicod and Kanavos 2012; Clement *et al.* 2009) and mixed methods approaches (Nicod and Kanavos 2016; Nicod 2017; Nicod *et al.* 2020). However, the difficulty of quantitatively analysing the topic of HTA coverage recommendations has led to a paucity of rigorously designed economic modelling studies, many of which are statistically underpowered and/or constrained to a single jurisdiction.

Some quantitative studies have emerged over time studying the factors that influence HTA coverage recommendations; for example, evidence uncertainty and disease burden were found to be important decision modifiers beyond incremental cost-effectiveness ratios (ICER) in the context of NICE (Devlin and Parkin 2004). Subsequent studies supported the assumption that ICER thresholds are not deterministic and that multiple factors can affect the probability of rejection (Dakin, Devlin, and Odeyemi 2006; Dakin *et al.* 2015). Similar single-country studies utilising logistic regressions exist for Scotland (Charokopou *et al.* 2015), Australia (Harris *et al.* 2008; Segal, Dalziel, and Mortimer 2010; Chim *et al.* 2010) and France (Le Pen, Priol, and Lilliu 2003).

The cross-country econometric literature is more limited. A study using survey data from 13 OECD countries to separately analyse the stages of evidence assessment and appraisal (i.e. decision-making), highlighted that evidence uncertainty has a significant influence on the likelihood of rejection, mitigated only by the participation of a greater number of stakeholders in the appraisal stage (Fischer *et al.* 2013). A study focusing on oncology drugs in the EU showed that macro-level health system characteristics do not account for differences in reimbursement of cancer drugs, apart from a significant positive association between SHI systems and coverage probability (Maynou and Cairns 2019). A further study has found that agencies adopt an extra-welfarist approach in decision-making by going beyond the clinical evidence and conventional

metrics of societal welfare, taking into account dimensions of value like drug innovativeness and unmet need (Kanavos, Visintin, and Gentilini 2023).

To date, there have been no published studies carrying out a comprehensive cross-country comparative analysis of HTA outcomes from the perspective of negative recommendations to understand what factor or combination of factors contribute to drugs being rejected by HTA agencies.

Against this background, therefore, this paper empirically studies the factors associated with HTA rejections (i.e., drugs receiving negative recommendations by HTA agencies) in a sample of prescription drugs assessed between 2009 to 2020 across eight HTA agencies in seven OECD countries. In doing so, we adopt a comparative approach to identify these factors. Additionally, we study the divergence in negative recommendations between agencies and explore agency-specific differences in the assessment of different types of drugs. Overall, we seek to answer two questions: first, what are the most significant factors contributing to HTA rejections across our sample? And, second, how does the importance of those factors differ across countries and agencies?

2. Data & methods

2.1. Sample selection and data collection

We extracted and used data from publicly available HTA reports. Eight HTA agencies across seven OECD countries were selected into the sample and fulfilled all four of the following criteria: first, they had established HTA processes operated by dedicated assessment bodies; second, they had diverse and explicitly stated decision-making criteria; third, they made HTA reports and information around recommendations publicly available; and, fourth, the information provided in the reports covered all aspects of assessment and appraisal to enable comparative analysis, particularly in what concerns the rationale for decision-making and the handling of clinical and economic uncertainties. The countries (and associated HTA agencies) that qualified were: Australia (PBAC – Pharmaceutical Benefits Advisory Committee); Canada (Canadian Drug Agency – CDA, formerly known as CADTH – Canadian Agency for Drugs and Technologies in Health, until 2023; and INESSS – Institut national d'excellence en santé et en services sociaux in Quebec); England (NICE – National Institute for Health and Care Excellence); France (HAS – Haute Autorité de Santé); Germany (IQWiG – Institute for Quality and Efficiency in Health Care); Scotland (SMC – Scottish Medicines Consortium); and Sweden (TLV – The Dental and Pharmaceutical Benefits Agency). Supplementary Table 1 reviews the regulatory structure and assessment methods followed by the sample HTA agencies.

Data extraction commenced with drugs that have received regulatory approval. As a first step, therefore, we considered all drugs that received marketing authorisation (MA) across the regulatory agencies in the study countries (European Medicines Agency - EMA for the European countries in the sample; Health Canada, for Canada and Quebec; and Therapeutic Goods Administration – TGA for Australia) between January 1st, 2009 and June 30th, 2018. In order to capture HTA outcomes for drugs that received MA in 2017 and 2018, we extended the study period to December 31st, 2020.

The selection process ensured that a representative sample of drugs and their respective indications ('drug-indication pairs') approved by regulatory authorities (EMA in Europe, TGA in Australia, and Health Canada in Canada) over the study period was created, by not restricting the sample only to drug-indication pairs assessed by all HTA agencies and including in the sample drug-indication pairs which were assessed by at least two HTA agencies in order to ensure comparability and minimise the sample's bias towards certain therapeutic areas. HTA reports that were published by the eight agencies between January 1st, 2009 to December 31st, 2020 were obtained by directly accessing the HTA agency websites. A spreadsheet dataset was constructed and populated with information found within the reports. We excluded from the sample generics and biosimilars, and, therefore, our sample consisted exclusively of originator prescription drugs.

The unit of analysis was defined as the 'drug-indication', which meant that HTA outcomes on different indications for the same drug were entered into the dataset separately. The dataset was structured so that each entry represented a single drug-indication-HTA agency triplet. This resulted in a single HTA decision per drug-indication and agency (or a missing observation if no HTA report was found by that agency for that drug-indication). This was achieved by applying the following rules when analysing HTA reports: first, the most up-to-date full submissions were selected for each drug-indication; second, for instances where a resubmission was made (e.g., following an earlier rejection), the outcome of the resubmission was recorded and not the outcome of any previous assessments; and, third, reports of minor amendments to previous recommendations, which did not impose any further important clinical or economic restrictions, were excluded from the analysis.

2.2. Dataset description

Our dataset is an unbalanced panel of 217 drug-indication pairs, representing 175 unique drugs (molecules), assessed by at least two out of the eight HTA agencies over the study period 2009-2020, resulting in a sample of N = 1,405 observations. No agency had appraised all 217 drug-indication pairs in our dataset. Overall, there were 49 drug-indication pairs (22.6%), which were assessed by all eight agencies over the study period, while a majority of 180 drug-indications (82.9%) were assessed by at least six out of the eight agencies (see Supplementary Table 2).

2.3. Variables of interest

Our outcome is a binary variable indicating a negative HTA outcome (=1 for "do not list" recommendation; =0 for "list or list with conditions" recommendation). The explanatory variables utilised in our analysis (Table 1) included the following: First, *regulatory variables* (defined to control for the pathway used in obtaining marketing authorisation and distinguishing between *standard marketing authorisation*, *conditional/accelerated marketing authorisation* and *parallel review*; and whether the drug-indication pair had a *previous rejection by the HTA agency*). It is hypothesised that a standard approval increases the odds of rejection compared with accelerated approval, since accelerated approval provides a signal that the drug in question holds significant promise of clinical benefit. Equally, it is hypothesised that a previous rejection by an HTA agency could increase the odds of a (final) rejection. Second, *disease-related variables* (classifying drugs by *cancer*, *non-cancer*, *orphan* and *non-orphan* indications), whereby the hypothesis to be tested was whether the use of a drug for a specific indication (cancer- vs non-cancer-related and orphan- vs non-orphan-related) increased the odds of a rejection by HTAs. Third, *clinical evidence variables* were included (controlling for the quality of evidence, using a composite index of quality drawing on similar research in earlier studies (Kanavos, Visintin, and Gentilini 2023; Sterne *et al.* 2019), and which was derived from information on (i) trial design (whether the evidence submitted was based on phase III Randomised Control Trial(s) (RCT) vs other form of trial design), (ii) comparator choice (direct comparator vs. placebo) and (iii) study primary endpoints utilised in the pivotal trial(s) submitted to the HTA agency (whether clinical vs surrogate). The hypothesis to be tested was whether poor quality submissions (e.g., relying on surrogate rather than clinical endpoints, or on phase II RCTs or single arm studies, rather than Phase III RCTs), increased the odds of a rejection by HTAs. Finally, *clinical and economic uncertainty variables* were included that could not be addressed by sponsors and which controlled for the perceived evidence uncertainties raised by HTA agency proceedings. Key clinical uncertainties included uncertainties relating to clinical benefit, comparator, practice generalisability, population generalisability, study design, follow-up period and adverse events. Key economic uncertainties included uncertainties relating to cost-effectiveness, utility input, cost input, economic modelling and economic model comparator. Each of these variables summarised

Table 1. Dataset variable descriptions and hypothesised effect on probability of rejection

Variable	Type*	Description	Hypothesised effect on odds of HTA rejection	
Regulatory Variables				
Marketing authorisation pathway	Cat	Pathway of approval prior to HTA submission	Standard approval: ↑ Accelerated approval: ↓ Parallel review: ↓	Early approvals and parallel reviews (only applicable in Australia, Canada, Quebec) are likely linked with higher quality submissions/ higher perceived importance of a drug
Previous HTA rejection	Bin	=1 if drug-indication previously rejected by agency	↑ or ↓	Outcome of submissions subsequent to a rejection depends on mitigating steps taken by the manufacturer in the interim (e.g. new evidence generation, addressing clinical or/and economic uncertainties)
Disease-related Variables				
Cancer indication	Bin	=1 if cancer drug	↑	Cancer drugs are expensive, often refer to small patient populations and have highly uncertain evidence submissions carrying a higher probability of rejection compared with non-cancer drugs
Orphan indication	Bin	=1 if drug-indication designated as orphan by at least one agency	↑ or ↓	Orphan drugs may be rejected at higher rates, but an orphan indication could also lower thresholds for accepting clinical evidence of lower quality (e.g. non-RCT data); countries may apply higher thresholds for listing orphan drugs; the effect, therefore, is unclear
Clinical Evidence Variables				
Quality of evidence	Cat	Composite index of quality derived from binary variables for (1) Phase III trial; (2) head-to-head trial; and (3) clinical endpoint	Low qual. (≤1/3): ↑ Average qual. (2/3): ~ High qual. (3/3): ↓	High-quality evidence means that the submission meets all 3 criteria; if the highest level of evidence submitted is lower in the evidence hierarchy, we expect a higher probability of rejection; therefore, average or low quality evidence carry higher odds of rejection
Phase III Trial	Bin	=1 if at least one phase III trial informing submission	↓	Phase III trials as opposed to phase II/I trials are a minimum accepted standard in most submissions; phase II or earlier phase trials or other trial designs carry higher odds of rejection
Head-to-head trial	Bin	=1 if at least one active comparator in trial informing submission	↓	Head-to-head trials using an active comparator over placebos are of higher relevance to decision makers; placebo-controlled trials carry higher odds of rejection
Clinical endpoint	Bin	=1 if at least one clinical endpoint in submitted trials	↓	Clinical endpoints measuring direct health outcomes (e.g. survival) over surrogate markers are expected to be of higher relevance to decision makers; surrogate markers carry higher odds of rejection

(Continued)

Table 1. (*Continued*)

Variable	Type*	Description	Hypothesised effect on odds of HTA rejection	
Regulatory Variables				
Not addressed Clinical Uncertainty Variables[†]				
Clinical benefit	Bin	=1 if at least one clinical benefit uncertainty	↑	Clinical benefit uncertainties relate to size of the (clinical) effect, where in many cases the benefit is considered inadequate; they are the most frequently raised concerns by HTA angecies and are likely to be the most significant amongst clinical evidence uncertainties, particularly for countries focusing on comparative clinical benefit (Germany, France); may require revision and submission of additional evidence or indirect treatment comparison (ITC)
Practice generalisability	Bin	=1 if at least one clinical practice generalisability uncertainty	↑	Concerns raised by HTA agencies around the generalisability of the evidence to local clinical practice are expected to have an effect but may be more easily overcome/ addressed by manufacturers with jurisdiction-specific evidence
Population generalisability	Bin	=1 if at least one population generalisability uncertainty	↑	Concerns raised by HTA agencies around the generalisability of study populations to the country context are infrequent but may be important when mentioned in HTA reports
Adverse events	Bin	=1 if at least one adverse events uncertainty	↑	Concerns raised by HTA agencies around clinical safety are important, particularly if they arise post-marketing authorisation
Study design	Bin	=1 if at least one study design uncertainty	↑	Study design uncertainties raised by HTA agencies are related to potential methodological gaps in clinical trial design (e.g., single-arm trials are very problematic for HTAs), which undermine the confidence in the evidence submitted
Long-term follow-up	Bin	=1 if at least one long-term effect (follow-up) uncertainty	↑	Follow-up concerns are likely to be related to study design and shorter patient follow-up periods, which decrease the confidence on submitted trial results
Comparator	Bin	=1 if at least one comparator uncertainty	↑	Comparator concerns raised by HTA agencies are related to both the type (e.g. placebo) and the appropriateness of the selected comparator in the clinical evidence submitted; acceptable comparators are linked to standard of care in the submitted setting rather than the comparator used in the clinical trial

(*Continued*)

Table 1. (*Continued*)

Variable	Type*	Description	Hypothesised effect on odds of HTA rejection	
Regulatory Variables				
Not addressed Economic Uncertainty Variables[†]				
Cost-effectiveness	Bin	=1 if at least one cost-effectiveness uncertainty	↑	Cost-effectiveness concerns raised by HTA agencies relate to possible under-estimates of submitted ICERs, which can be due to clinical benefit over-estimates, eligible cost under-estimates or utility value concerns (or a combination of these) in settings that use cost-effectiveness as a criterion for coverage and can lead to rejection or request for revision by sponsor
Utility input	Bin	=1 if at least one utility (in economic model) uncertainty	↑	Utility input concerns raised by HTA agencies relate to uncertainties in the calculation of the utility metrics utilised in the model; they may lead to rejections and request for revision by sponsor
Cost input	Bin	=1 if at least one cost (in economic model) uncertainty	↑	Cost input concerns relate to the cost figures utilised in the economic model leading to cost under-estimates; they may lead to rejections and request for revision by sponsor
Modelling	Bin	=1 if at least one modelling/model type uncertainty	↑	Concerns raised by HTA agencies surrounding the validity of an economic model (e.g., inappropriate model or modelling approach) are frequent and are likely to lead to rejections and request for revision by sponsor
Comparator input	Bin	=1 if at least one comparator choice uncertainty (in the economic model)	↑	Comparator concerns raised by HTA agencies are related to both the type (e.g. placebo) and the appropriateness of the selected comparator in the economic model submitted and can lead to rejections and/or request for revision by sponsor
Fixed Effects				
HTA agency	Cat	Captures agency-specific time-invariant characteristics	–	–
Assessment year	Cat	Captures year-specific agency-invariant influences	–	–

Notes: *Bin = binary (YES = 1, NO = 0) variable; Cat = categorical variable.
[†]An uncertainty was categorised as 'unaddressed' if **none** of the following criteria were fulfilled: A) The manufacturer presented more evidence; B) External stakeholders (e.g. professional medical bodies) interfered to address uncertainty; C) The HTA agency dismissed the uncertainty as unimportant.

concerns raised by HTA agencies around the extent to which the submitted clinical or economic evidence fulfilled optimal or acceptable thresholds. The hypothesis to be tested here was that submissions to HTA agencies with significant uncertainties that could not be addressed by sponsors carried a higher risk of rejection by HTAs.

The data and variables informing our analysis have been extracted manually by two of the authors (PK and EV) from the extracted HTA reports and were cross-validated by all authors. Specifically, for clinical and economic uncertainties, manual extraction involved the authors identifying the relevant concerns raised by HTA agencies in the appraisal part of the HTA report, which informed the final recommendation. Table 1 summarises the variables, their definitions, the hypotheses they are testing and the expected or likely direction of the effect.

2.4. Empirical methods

A descriptive analysis preceded our econometric analysis and guided our selection of independent variables and model specification. In all our models, we employ multivariate logistic regressions, with a binary outcome capturing the likelihood of a negative HTA outcome for a drug-indication by an agency. Our models controlled for a comprehensive set of independent variables, as defined in Table 1. First, we accounted for assessment-specific characteristics, which vary at both drug-indication and agency levels (regulatory pathway, quality of evidence, previous rejection, clinical and economic uncertainties). Second, we included agency-invariant controls, which change only across drug-indications, such as disease-related variables (cancer vs non-cancer drugs; orphan vs non-orphan drugs). Last, we also controlled for agency fixed-effects to account for time-invariant differences across agencies, and year fixed effects to account for year-specific variation(s). We further examined whether the association between drug-indication characteristics and HTA decisions varied across agencies, by including interaction terms in our regression models (e.g., cancer drugs with orphan indications). The strength of associations are quantified using odds ratios and marginal effects (MEs), which allow us to illustrate how the predicted probability of a negative HTA outcome changes with potential changes in the independent variables presented above (Norton et al, 2019). Further detail about the model specification can be found in the Supplementary Material.

All analyses were carried out in STATA 16.1 ('StataCorp. 2019. Stata Statistical Software: Release 16. College Station, TX: StataCorp LLC.'). Two sets of results are shown, one for the full sample of 8 HTA agencies and one focusing only on countries pursuing cost-effectiveness analysis, therefore, excluding France and Germany, which are pursuing HTA with comparative clinical benefit assessment. Estimates are presented in the form of odds ratios (OR). When examining inter-agency differences, we additionally compute average MEs to aid the interpretation of interactions in our models. Standard errors clustered at the agency level are reported throughout.

3. Results

3.1. Descriptive analysis

Out of a total of 1,405 HTA assessments, there were 181 'do not list' (reject) HTA outcomes by the end of the study period (Table 2). Rejections represented 12.9% of all assessments, largely because only final decisions were recorded even if that drug-indication had been previously rejected by the agency. There were 281 HTA assessments where agencies had previously rejected a drug-indication but subsequently issued a positive coverage recommendation ("list" or "list with criteria").

Of the 217 drug-indications assessed by at least two agencies, there were 110 drug-indications (50.7%) which at least one agency rejected while some or all others recommended it for reimbursement, suggesting a significant divergence in HTA outcomes. Out of the 110 drug-indications rejected, most were rejected by less than half of the agencies that assessed them (n = 93, 84.5%). Only a single drug-indication pair was rejected (Nivolumab for treating locally advanced unresectable or metastatic urothelial cancer after platinum-containing chemotherapy) by both agencies that assessed it, namely, NICE and SMC (Supplementary Table 3).

Table 2. Descriptive statistics for variables employed in the econometric models, presented by HTA outcome (rejection vs. listing)

Variable Count (% of column total [n])	Rejections (DNL)* n = 181	Listings (L/LWC)* n = 1,224	All Assessments N = 1,405
Regulatory Variables			
Marketing authorisation pathway | $x^2 = 1.68, p = 0.431$			
– Standard approval	128 (71%)	918 (75%)	1,046 (75%)
– Early Approval	33 (18%)	182 (15%)	215 (15%)
– Parallel Review	20 (11%)	124 (10%)	144 (10%)
Previous HTA rejection | $t = 2.33, p = 0.010$	57 (31%)	281 (23%)	338 (24%)
Disease-related Variables			
Cancer indication | $t = 1.80, p = 0.036$	77 (43%)	434 (35%)	511 (36%)
Orphan indication | $t = 2.39, p = 0.009$	49 (27%)	229 (19%)	278 (20%)
Cancer & orphan indication | $t = 0.36, p = 0.361$	23 (13%)	144 (12%)	167 (12%)
Non-cancer & non-orphan indication | $t = -3.67, p < 0.001$	78 (43%)	705 (58%)	783 (56%)
Clinical Evidence Variables			
Quality of evidence | $x^2 = 21.29, p < 0.001$			
– Low quality	104 (57%)	501 (41%)	605 (43%)
– Average Quality	56 (31%)	599 (49%)	655 (47%)
– High Quality	21 (12%)	124 (10%)	145 (10%)
Phase III trial | $t = -1.51, p = 0.067$	149 (82%)	1,063 (87%)	1,212 (86%)
Head-to-head trial | $t = -4.81, p < 0.001$	67 (37%)	681 (56%)	748 (53%)
Clinical endpoint | $t = 1.92, p = 0.028$	42 (23%)	206 (17%)	248 (18%)
Not addressed Clinical Uncertainty Variables[†]			
Clinical benefit | $t = 4.39, p < 0.001$	92 (51%)	409 (33%)	501 (36%)
Practice generalisability | $t = 2.09, p = 0.019$	42 (23%)	199 (16%)	241 (17%)
Population generalisability | $t = 1.00, p = 0.158$	27 (15%)	148 (12%)	175 (12%)
Adverse events | $t = 2.12, p = 0.018$	39 (22%)	180 (15%)	219 (16%)
Study design | $t = 2.21, p = 0.014$	70 (39%)	369 (30%)	439 (31%)
Follow-up | $t = 1.02, p = 0.155$	76 (42%)	465 (38%)	541 (39%)
Comparator | $t = 0.55, p = 0.292$	21 (12%)	125 (10%)	146 (10%)

(*Continued*)

Table 2. (*Continued*)

Variable Count (% of column total [n])	Rejections (DNL)* n = 181	Listings (L/LWC)* n = 1,224	All Assessments N = 1,405
Restricted sample of CEA countries¶	Rejection (DNL) n = 170	Listing (L/LWC) n = 887	All Assessments N = 1,057
Not addressed Economic Uncertainty Variables†			
Modelling │ *t* = 0.06, *p* = 0.477	60 (35%)	311 (35%)	371 (35%)
Cost input │ *t* = 1.46, *p* = 0.07	42 (25%)	173 (20%)	215 (20%)
Utility input │ *t* = 3.04, *p* = 0.001	38 (22%)	107 (12%)	145 (14%)
Comparator input │ *t* = 1.27, *p* = 0.103	19 (11%)	70 (8%)	89 (8%)
Cost-effectiveness │ *t* = 4.00, *p* <0.001	72 (42%)	231 (26%)	303 (29%)

Notes: Chi-square and t-statistics reported compare sample of rejections with sample of listings.
*L = recommend drug for listing within the marketing authorisation (MA) indication; LWC = recommend drug conditional on indication restriction (population, dosage, cycles of therapy, etc.); DNL = do not recommend drug for reimbursement.
¶CEA = cost-effectiveness assessment; includes all countries & agencies apart from France (HAS) and Germany (IQWiG).
†An uncertainty was categorised as 'unaddressed' if **none** of the following criteria were fulfilled: A) The manufacturer presented more evidence; B) External stakeholders (e.g. professional medical bodies) interfered to address uncertainties; C) The HTA agency dismissed the uncertainty as unimportant.

Examining the pathways that manufacturers used prior to HTA submission, we find that the 'early approval' pathway, known as conditional marketing authorisation or accelerated approval, seems to be sought in submissions with lower quality of evidence, reflecting the early stage of the trial data and/or the difficulty of conducting high-quality studies due to restricted patient population sizes, as is often the case in most orphan indications (Figure 1). The drug-indications that are eventually rejected seem to make greater use of the early approval pathway (18% of 'rejected' drug-indications apply for 'early approval' vs 15% of those that have achieved 'listed' or 'listed with conditions') and have lower quality evidence associated with these submissions compared to listed drugs (82% of rejected submissions following an 'early approval' have low quality evidence vs 47% of listed); the higher proportion of low-quality evidence in rejected drugs versus listed drugs in early approvals is statistically significant (t= −3.76, p < 0.001). By contrast, the parallel review pathway seems to be used when evidence of high or average quality is subsequently submitted to the HTA regulator (60% of 'rejected' and 66% of 'listed' submissions assessed via parallel review had evidence of high or average quality), possibly reflecting the manufacturer's confidence in the submission being accepted first-time round.

A comparative analysis of rejection rates across country agencies shows several trends (Figure 2). First, rejection frequency across agencies varies, with Quebec and Australia having rejected the most submissions at 36% and 20%, respectively (Figure 2a). We also note that Germany and France have particularly low rejection rates when compared to the other countries (because only a single drug-indication in our sample has been rejected in Germany – Olaparib for ovarian cancer – we exclude Germany when examining inter-agency differences). This is in line with their different HTA approach that focuses solely on comparative clinical benefit assessment and the particularities of their pricing and reimbursement system, which allow drugs to be subjected to internal reference pricing when their comparative clinical benefit is found to be non-important, instead of being outright rejected.

Second, we observe that drugs with a cancer and/or orphan indication represent a significant proportion of rejections across all agencies, with England's NICE rejections being skewed most
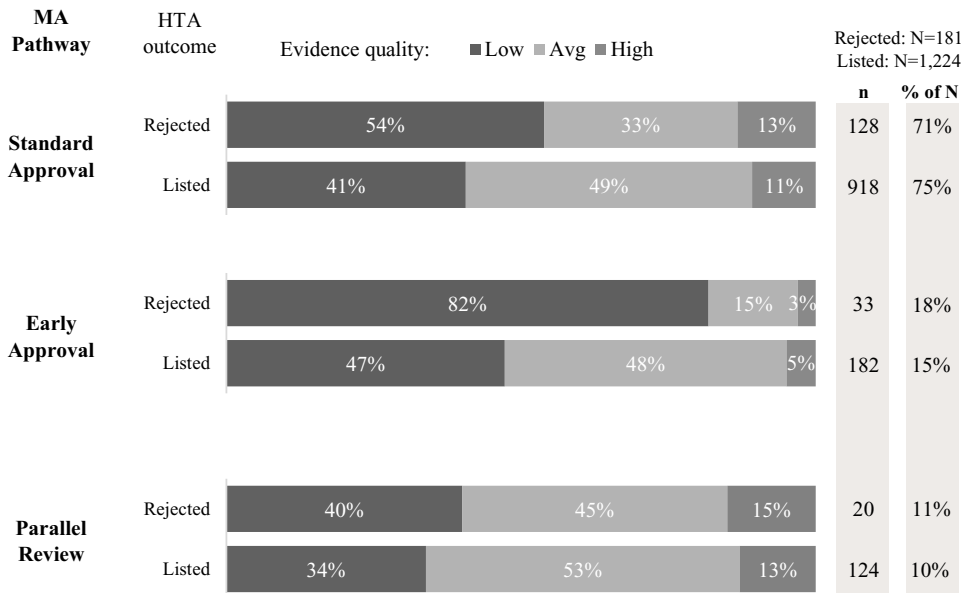
Figure 1. Breakdown of HTA submissions by (a) marketing authorisation (MA) pathway, (b) quality of evidence (low, average, and high), and (c) HTA outcome (rejected vs listed drug-indications)[1]. *Note:* [1]Each of the evidence quality bars represents the % of all submissions (n) that utilised a particular MA pathway and had one of the two HTA outcomes.

heavily towards cancer drugs despite the agency rejecting a below-average number of the cancer drugs it assessed over the study period (14% of assessed cancer drugs were rejected by NICE vs. sample average of 17%). On that point, it is interesting to note that one-third of the 80 cancer drugs in the sample also have an orphan indication (n = 26, 33%), which reflects the 'personalised medicine' shift over the past decade that has brought highly-specific and targeted cancer therapies to market.

Third, examining rejection trends over time reveals that Quebec, Australia and Scotland have, in general, rejected a greater proportion of submitted drugs from 2016 onwards (Figure 2b).

Finally, both listed and rejected drug-indication pairs present clinical and economic uncertainties, but a significantly higher proportion of rejected drugs is linked to both clinical and economic uncertainties; for example, 57% of rejections are linked to clinical benefit uncertainties compared with 33% of listings, and the same holds for practice generalisability (23% vs 16%), population generalisability (15% vs 12%), adverse events (22% vs 15%), inadequate follow up (42% vs 38%) and comparator concerns (12% vs 10%). Similar differences exist in the context of economic uncertainties, for example, in 42% of all rejections cost-effectiveness concerns were raised compared with 26% in all listing recommendations, and similar was the case in utility inputs (22% vs 12%), cost inputs (25% vs 20%) and economic model comparator inputs (11% vs 8%).

3.2. Factors associated with HTA rejections

In the model including all study countries (full sample), we used two alternative covariate specifications, the first (Model [1]) controlling for regulatory, disease, submission and clinical evidence characteristics, while the second (Model [2]) additionally controlling for non-addressed clinical evidence uncertainties; models [3] and [4] replicate the analysis for a restricted sample of countries using cost-effectiveness analysis ('CEA countries') for resource allocation purposes, therefore excluding France and Germany (the countries using comparative clinical benefit as the

(a)   Distribution of HTA rejections by indication, % of all HTA rejections, 2009-2020.

■ Cancer indication only ■ Orphan indication only ▨ Cancer & orphan indication ▨ Other Drugs

	England	Scotland	Sweden	France	Australia	Canada	Quebec
Rejections	11	18	10	10	37	29	65
% of all assessments rejected	7%	9%	7%	5%	20%	15%	36%
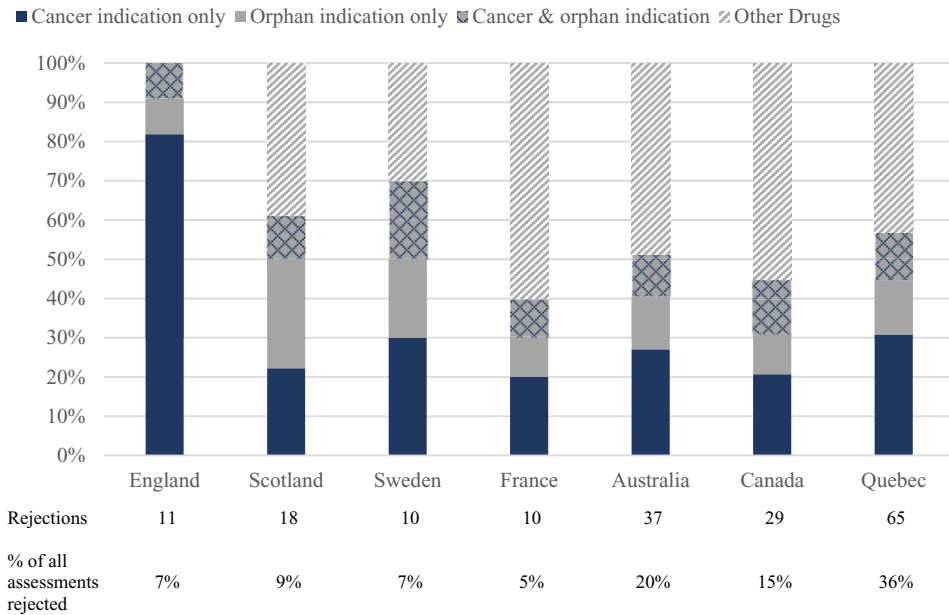
(b)   Rejection rate over time by agency, 3-year moving average, % of submissions rejected, based on year of marketing authorisation (2009-2018)
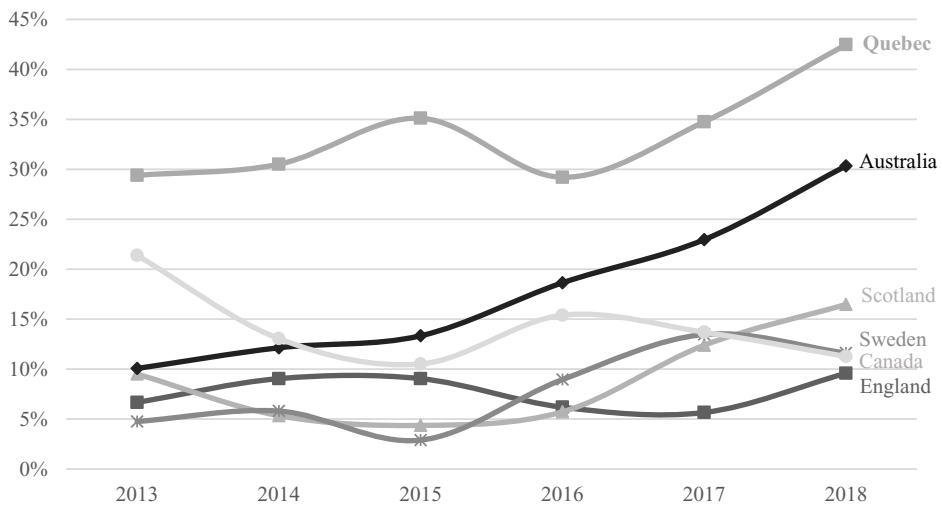
Figure 2. Rejection rates across countries (a) by drug indication and (b) over time. *Notes:* An indication is recognised as orphan if any agency in the sample designated that drug-indication as an orphan one. This methodology is adopted because definitions of a rare disease may vary and some countries do not formally recognise orphan status in HTA proceedings (Australia, Canada, Quebec). Germany is excluded from this analysis because it rejected only a single drug (Olaparib) in the study period.

main decision criterion) (Table 3). In the restricted sample, we additionally control for whether an economic model was submitted, and whether a set of unaddressed economic evidence uncertainties were reported. All regressions control for agency and assessment year fixed effects (FE). Sensitivity analyses were conducted by including and excluding independent variables in the regression models to ensure robustness and absence of multicollinearity. Finally, a detailed look into the impact of unaddressed clinical and economic uncertainties on odds of rejection can be found in Supplementary Table 4.

In the full sample models, of the three regulatory pathways, only a parallel review process seems to be consistently associated with a lower probability of rejection (OR ranging between 0.43 and 0.41, 95%, $p < 0.01$). Interestingly, this relationship is driven by only three agencies that allow manufacturers to submit a parallel review request (Canada, Australia and Quebec), which are also the agencies with the highest overall rejection rates (15%, 20% and 36%, respectively). We can link this finding to our original hypothesis that drugs with evidence of higher quality for which manufacturers are confident about their approval by HTA agencies utilise the parallel review pathway to achieve faster market access. This relationship holds even when controlling for the quality of evidence and reported unaddressed evidence uncertainties, signalling a favourable treatment of the drugs that utilise this pathway over and above what the quality of evidence alone would predict. For the early approval pathway, we find a similar reduction in the odds of rejection as with parallel reviews but the strength and statistical significance of the association are less compelling.

In non-cancer drugs, an orphan indication is associated with a strong increase in the odds of rejection (OR ranges between 2.30 and 2.37, $p < 0.01$). Similarly, a cancer indication is associated with higher odds of rejection (model [1], OR 2.17, $p < 0.05$, model [1]), although this diminishes in significance when controlling for the likely poorer clinical evidence and higher uncertainties in their submissions (model [2], OR 1.76, where the effect is not significant at the 5% level). However, when considering the interaction term, it appears that the combination of an orphan and cancer indication considerably reduces the likelihood of rejection (OR ranges between 0.28 - 0.26, $p < 0.01$). Overall, an orphan indication seems to be a more powerful predictor of a negative outcome for both cancer and non-cancer drugs.

The impact of quality of evidence on rejection probability is significant when considering the composite index made up of the type of trial (at least one Phase III), the type of comparator (at least one active comparator/head-to-head trial) and the type of endpoints (at least one clinical endpoint). Low-quality evidence expectedly increases the odds of rejection (OR ranging between 2.07, in model [1], $p < 0.01$; and 2.20, in model [2], $p < 0.01$). When considering the impact of unaddressed clinical uncertainties on the odds of rejection, only concerns around clinical benefit increase the odds of rejection (model [2], OR 2.59, $p < 0.01$). Finally, a previous HTA rejection does not seem to change the odds of rejection vs listing and the effect is not statistically significant.

The above results are very similar to those obtained in the restricted sample comprising cost-effectiveness countries only ('CEA countries). For the CEA countries only, the use of economic modelling in submissions is protective against a negative recommendation, although virtually all submissions to CEA countries make use of an economic model (96%). Uncertainties around cost-effectiveness has an important effect on increasing the probability of rejection (model [4], OR 2.07, $p < 0.01$) and the same holds for utility evidence used in economic models (model [4], OR 2.38, $p < 0.01$).

3.3. Inter-agency differences

The observed divergence in decision-making is assessed by, first, exploring differences in overall rejection rates and, second, examining the differences in the effect of selected regressors by agency. Canada, Quebec and Australia have considerably higher overall rejection rates compared to England, while France expectedly has lower rejection rates. Surprisingly, the closely related

Table 3. Logistic regressions of HTA rejection for the full sample and a restricted sample of countries using cost-effectiveness analyses in assessments which excludes France (HAS) and Germany (IQWiG)

Coefficient Format: Odds Ratios (Clustered SE)	Full Sample (8 HTA agencies)		CEA settings (6 HTA agencies)*	
	Model [1]	Model [2]	Model [3]	Model [4]
Marketing authorisation pathway (vs. standard approval)				
Early approval	0.709	0.610**	0.709	0.637
	(0.148)	(0.129)	(0.169)	(0.167)
Parallel review	0.434***	0.408***	0.453***	0.458***
	(0.077)	(0.064)	(0.073)	(0.075)
Drug indication				
Orphan indication (vs non-orphan indication)	2.301***	2.368***	2.680***	2.574***
	(0.693)	(0.706)	(0.963)	(0.943)
Cancer indication (vs non-cancer indication)	2.167**	1.758*	2.691**	1.736
	(0.734)	(0.593)	(1.197)	(0.694)
Cancer with orphan indication (vs all other drugs in sample)	0.280***	0.258***	0.210***	0.234***
	(0.126)	(0.118)	(0.113)	(0.107)
Previous HTA rejection (y/n)	0.937	0.883	0.953	0.922
	(0.098)	(0.101)	(0.111)	(0.174)
Low quality of evidence (y/n)	2.069***	2.198***	1.764***	1.874***
	(0.436)	(0.472)	(0.340)	(0.389)
Clinical benefit uncertainty (y/n)		2.587***		2.306***
		(0.352)		(0.355)
Economic model submitted (y/n)				0.125
				(0.139)
Cost-effectiveness uncertainty (y/n)				2.075***
				(0.581)
Utility model input uncertainty (y/n)				2.384**
				(0.886)
Number of obs.	1,395	1,395	1,048	1,048
Pseudo R^2	0.178	0.210	0.164	0.222
AIC	899.4	865.0	784.5	730.7
BIC	936.1	901.7	809.2	755.5
Assessment Year FE	Yes	Yes	Yes	Yes
Agency FE	Yes	Yes	Yes	Yes
Unaddressed clinical uncertainties	No	Yes	No	Yes
Unaddressed economic uncertainties	No	No	No	Yes

Notes: *CEA = cost-effectiveness assessment; includes all countries & agencies apart from France (HAS) and Germany (IQWiG).
Cluster robust SE in parentheses; Significance levels: ***$p < 0.01$, **$p < 0.05$.

Table 4. Average marginal effects (ME) of interactions of selected variables with agency dummies[1]

Average Marginal Effects (Interactions with *agency*) - Model [2]					
dy/dx	Cancer indication	Orphan indication	Cancer & orphan indication	Clinical benefit uncertainty	Cost-effectiveness uncertainty
England (NICE)	0.0930***	–0.0412***	–0.103***	0.0159	0.110***
	(0.0118)	(0.0134)	(0.0128)	(0.0176)	(0.0162)
Scotland (SMC)	0.0122	0.166***	–0.0703**	0.130***	0.118**
	(0.0358)	(0.0271)	(0.0276)	(0.0273)	(0.0525)
Sweden (TLV)	0.0820***	0.101***	–0.0269	0.113***	0.00358
	(0.0184)	(0.0236)	(0.0292)	(0.0215)	(0.00553)
France (HAS)[†]	–0.0233***	–0.0125	–0.0483***	0.0380***	–
	(0.00709)	(0.0117)	(0.0103)	(0.0102)	–
Australia (PBAC)	–0.0936***	–0.0128	–0.174***	0.0544	0.0603
	(0.0203)	(0.0187)	(0.0248)	(0.0410)	(0.0466)
Canada (CDA)	–0.00603	0.0542***	–0.0848**	0.0841***	0.0149
	(0.0203)	(0.0185)	(0.0333)	(0.0277)	(0.0238)
Québec (INESSS)	0.0931***	0.0170	–0.215***	0.170**	0.202***
	(0.0351)	(0.0536)	(0.0475)	(0.0743)	(0.0592)
Number of obs. (France)[†]	1,301	1,259	1,256	1,307	1,388
Number of obs. (all others)	1,048	1,048	1,048	1,048	1,048

Notes: 1 Each column represents a logistic regression of HTA rejection after interacting the respective regressor with agency dummies, controlling for covariates specified in model [2];
†Marginal effects for France (HAS) are reported from the full sample model [2] which excludes unaddressed economic uncertainty dummies. For all other countries, the CEA country restricted model [2] is used, which includes unaddressed economic uncertainty dummies.
Cluster robust SE in parentheses; Significance levels: ***$p < 0.01$, **$p < 0.05$.

agencies of Scotland and Sweden also seem to exhibit higher rejection rates compared to England's NICE, although these apparent differences become insignificant when accounting for the effect of unaddressed economic evidence uncertainty (particularly cost-effectiveness concerns) in the restricted sample of CEA countries (Supplementary Table 5).

Examining the differences in the effect of selected predictors by agency in Table 4, reveals a mixed picture. The findings suggest that England, Sweden and Quebec are more likely to reject cancer drugs (9.3%, 8.2% and 9.3%, higher probability of rejection compared to non-cancer drugs, respectively), while Australia and France are more likely to approve cancer drugs (9.4% and 2.3% reduction in the probability of rejection, respectively). Sweden is more likely to reject orphan drugs as well (10.1% higher probability of rejection compared to non-orphan drugs), whereas England seems more likely to approve these (4.1% lower probability of rejection). Scotland and Canada are also more likely to reject drugs with orphan indications (16.6% and 5.4% higher probability of rejection, respectively). However, for cancer drugs with an orphan indication, we find a strong universal reduction in the probability of rejection across all agencies (2.7% [Sweden] to 21.5% [Quebec] reduction in the probability of rejection across the study countries compared to non-cancer, non-orphan drugs). All agencies have a higher propensity to reject drugs when there are unaddressed clinical benefit uncertainties, the effect being most prominent in Quebec (17% increase in probability of rejection compared to submissions with no clinical benefit

uncertainties). The same applies for unaddressed cost-effectiveness uncertainties, although a significant effect is found only in England (11% increase in rejection probability), Scotland (12% increase) and Quebec (20% increase).

4. Discussion

Our empirical approach presents a novelty which extends beyond simple descriptive analyses of HTA rejections. This is because controlling for several important confounders simultaneously, allows us to make inferences about the systematic differences in drug assessments across and between HTA agencies in our sample. There are several important points which should be taken away from our results.

First, while our research largely confirms earlier findings on the important drivers of HTA decision-making, it goes beyond to highlight that higher quality evidence in the form of head-to-head Phase III trials, which utilise direct clinical endpoints, have fewer clinical benefit uncertainties and higher cost-effectiveness all act protectively against a 'do not list' recommendation. Our analysis of cross-country differences in rejection rates confirms that important differences exist in the assessment of the same evidence by different agencies, even after accounting for the regulatory individualities of each. Such differences may be due to differences in the interpretation of evidence but also the existence of differences in social value judgements, such as rarity, severity, unmet medical need, or impact on family, among others, which may be explicitly incorporated into decision-making, or implicitly recognised on a per case basis at appraisal level.

Second, we find that, *ceteris paribus,* an orphan indication significantly reduces the risk of rejection in cancer drugs compared with non-cancer drugs. There are several challenges associated with the assessment of cancer medicines, which complicate decision-making, including high degrees of uncertainty, particularly in the size of clinical benefit, the type of endpoint used, and generalisability of the evidence and study populations, as there are often multiple indications per product, with some being combination therapies. However, an orphan indication in oncology medicines seems to sufficiently overcome the evidence shortfalls that may otherwise exist, reducing the likelihood of rejection. Presumably, the smaller expected budgetary impact of orphan oncology drugs (where the indication is narrower) and the higher level of unmet need makes them less prone to rejection compared to non-orphan cancer drugs.

Third, the consistency of our findings with respect to the association between orphan indications, cancer drugs and rejection probability may be misleading if we do not consider the 'early access' funding arrangements that are often in place for such treatments. For example, the UK's Cancer Drugs Fund, which operated separately from NICE until 2016, has provided funding for many promising oncology therapies which had not demonstrated sufficient cost-effectiveness at the time of their initial assessment (Wood and Hughes 2020). Similar earmarked funds and compassionate use programmes for highly-specialised medicines have operated throughout other OECD countries (e.g. the French ATU, provincial cancer agencies in Canada) (Degrassat-Théas *et al.* 2013; Srikanthan *et al.* 2018). The existence of such 'safety nets' for access to important therapies which are rejected by HTA agencies may be creating a propensity for rejecting drugs which can be made available to patients through alternative funding avenues, particularly when considering the non-orphan oncology drugs. We should not also discount the uptake and use of financial risk-sharing agreements as an integral part of coverage negotiations across OECD countries (OECD 2020).

Fourth, we illustrate that certain process-related differences across agencies are non-trivial. Significant inter-dependencies in HTA practices seem to exist for closely linked health systems such as those in the UK and Sweden, as well as in France and Germany. HTA rejection rates are generally low for the aforementioned countries but are quite high in Canada and Australia, statistically powering the samples of the latter countries to reveal some important trends. The lack of an explicit

orphan drug assessment framework in Canada (McMillan and Campbell 2017) during our study period has likely led to our finding of a considerably increased rejection rate for orphan drugs by CDA. Presumably, if some of these drugs were formally designated as orphan, the impact of the orphan indication on rejection probability would have been lower. On the contrary, the establishment of the pan-Canadian Oncology Drug Review (pCODR) in the Canadian context in 2010 does seem to have led to non-biased reviews of cancer drugs, as the lack of a similar such review pathway by the Canadian province of Quebec seems to be driving a strong positive association between oncology indications and rejections in that province. In Australia, there is a similar lack of an orphan drug assessment framework as in Canada, but a separate funding arrangement for orphan drugs exists in the form of the Australian Life Saving Drugs Program (LSDP). The LSDP seems to have had a positive impact on the assessments of orphan drugs, reducing the likelihood of their rejection in contrast to the situation in Canada. The LSDP provides patients with access to drugs treating life-threatening rare diseases but strict listing criteria has meant that few drugs have been included in the LSDP list since its inception in 1995 (Maguire and Tilden 2013; Australian Government Department of Health 2020).

4.1. Policy implications

A number of policy-relevant conclusions can be drawn from our analysis. First, the finding that certain agencies do not have processes tailored to the assessment of highly specialised or personalised treatments points to a need for reforming review pathways to consider drugs with atypical submissions separately. This is particularly important as an increasing number of highly individualised treatments are coming to market, rendering the conventional requirements of HTA agencies for randomised-controlled trial (RCT) evidence and robust cost-effectiveness analyses unrealistic. Indirect treatment comparisons and real-world evidence generation are obvious ways of overcoming evidentiary inadequacies, and are often combined with risk mitigation strategies, but their acceptability is often variable across settings. While randomised controlled trials have an important place in generating evidence, they are sometimes impossible as, for example, in the case of very rare diseases; and the lack of equipoise may render them unethical. Alternative trial designs, may also need to be implemented more broadly, particularly in the context of ever-increasing reliance on personalised medicine. These newer, or less familiar, study designs include basket trials, umbrella trials, Mendelian randomisation trials, step-wedge trials, ring trials, Baskerville trials and micro-randomised trials. HTA agencies will have to learn to use these newer designs if they are to remain relevant to evaluating evidence and to the practice of medicine and public health (Rawlins, 2016). Although this may be an issue for the future, current pharmaceutical innovation may expedite this process.

Second, the widespread use of 'early patient access' schemes may be inadvertently affecting decision-making at the HTA level, relaxing requirements for cost-effectiveness and making it easier for manufacturers to enter markets prematurely. On that note, we argue that risk-mitigation strategies will be vital to ensure that only innovation offering significant additional clinical benefits to patients is rewarded and perverse incentives for skipping regulatory hurdles are not created.

Third, the persisting divergence in coverage recommendations across agencies strengthens the argument for increasing cooperation in HTA decision-making and reinforcing the role of international collaborative initiatives in sharing HTA 'best practices' (e.g. INAHTA, EUnetHTA), at least in the context of having some common denominators vis-à-vis the assessment of new drugs and dealing with uncertainty. The EU regulation on HTA cooperation is one step in that direction, at least in the European context, provided a common ground is reached on the complexity of requirements to suit different Member States' needs (Kanavos et al., 2019; Main et al, 2025).

Finally, the association of parallel review processes with a lower probability of rejection needs to be linked to the fact that drugs with high-quality evidence stand better chances of being approved by HTA agencies. Hence, parallel review pathways should be more widely available in

order to expedite access to drugs that come to market with high-quality evidence and uncertainties that can be addressed by sponsors in a reasonable timeframe. Outstanding clinical evidence uncertainties can be addressed in a variety of ways, including evidence development in real-world settings and pay-for-performance schemes.

4.2. Study limitations

Our study has a number of limitations. The data collection methods which relied on analysts coding information found in published sources inevitably entail a certain degree of bias, partially mitigated by having multiple analysts review the same HTA reports and resolve conflicts through discussion. The absence of quantitative data on cost-effectiveness ratios and budget impact are also severely limiting our ability to draw meaningful conclusions about the results of our models. However, the frequent absence of these variables due to price confidentiality are a challenge for all such HTA studies. The modelling approach we employ allows us to raise some important points about the agencies in our sample, but the external validity of our conclusions remains unavoidably limited. Despite the considerable number of rejections which were analysed in our study compared to previous published papers, the rarity of the rejection event for some of the agencies leads to imprecisely calculated estimates and restricts statistical inference. Furthermore, a mixed methods approach using qualitative data from surveys or interviews with decision-makers may have been more appropriate to uncover the underlying mechanisms behind some of our findings. Future analyses with a larger sample would also enable a hierarchical model to be applied. Finally, the study findings pertain to specific HTA agencies within a defined time period, during which particular drug-indication pairs were assessed. Therefore, they should be interpreted with caution and should not be generalised to other contexts.

5. Conclusion

Having studied a sample of 1,405 drug-indication pairs, we have found that only 12.9% of the drugs are rejected by the 8 HTA agencies in our sample. Failure to satisfy HTA agencies is due to several factors relating to a product's indication (cancer or orphan indication not seen in combination), the quality of evidence (low quality evidence in the form of a placebo-controlled, early phase trials with surrogate endpoints) and the existence of clinical benefit and cost-effectiveness uncertainties raised during the assessment process and which remain unaddressed by sponsors. We have demonstrated that persisting differences exist in the stringency of assessments carried out by different HTA agencies for different subsets of drugs, highlighting the importance of considering contextual factors when understanding the reasons behind reimbursement decisions. Future research will aim to examine drug assessments longitudinally, to determine how drug reimbursement decisions can change over time and analyse the implications of delays caused by multiple assessments on timely patient access to life-saving treatments.

References

Akehurst RL, Abadie E, Renaudin N, and Sarkozy F. (2017) Variation in health technology assessment and reimbursement processes in Europe. *Value in Health* **20**(1), 67–76. doi: 10.1016/j.jval.2016.08.725.

Allen N, Liberti L, Walker SR, and Salek S. (2017) A comparison of reimbursement recommendations by European HTA agencies: is there opportunity for further alignment? *Frontiers in Pharmacology* **8**(June), 384. doi: 10.3389/fphar.2017.00384.

Allen N, Pichler F, Wang T, Patel S, and Salek S. (2013) Development of archetypes for non-ranking classification and comparison of European national health technology assessment systems. *Health Policy* **113**(3), 305–312. doi: 10.1016/j.healthpol.2013.09.007.

Angelis A, Lange A, and Kanavos P. (2018) Using health technology assessment to assess the value of new medicines: results of a systematic review and expert consultation across Eight European countries. *The European Journal of Health Economics: HEPAC: Health Economics in Prevention and Care* **19**(1), 123–152. doi: 10.1007/s10198-017-0871-0.

Australian Government Department of Health. (2020) 'About the Life Saving Drugs Program'. Text. Australian Government Department of Health. Australian Government Department of Health. October 9. https://www.health.gov.au/initiatives-and-programs/life-saving-drugs-program/about-the-lsdp.

Beletsi A, Koutrafouri V, Karampli E, and Pavi E. (2018) Comparing use of health technology assessment in pharmaceutical policy among earlier and more recent adopters in the European Union. *Value in Health Regional Issues* **16**(September), 81–91. doi: 10.1016/j.vhri.2018.08.002.

Charokopou M, Majer IM, Raad Jd, Broekhuizen S, Postma M, and Heeg B. (2015) Which factors enhance positive drug reimbursement recommendation in Scotland? A retrospective analysis 2006–2013. *Value in Health* **18**(2), 284–291. doi: 10.1016/j.jval.2014.12.008.

Chim L, Kelly PJ, Salkeld G, and Stockler MR. (2010) Are cancer drugs less likely to be recommended for listing by the pharmaceutical benefits advisory committee in Australia?' *PharmacoEconomics* **28**(6), 463–475. doi: 10.2165/11533000-000000000-00000.

Clement FM, Harris A, Jing Li J, Yong K, Lee K M, and Manns BJ. (2009) Using effectiveness and cost-effectiveness to make drug coverage decisions: a comparison of Britain, Australia, and Canada. *JAMA* **302**(13), 1437–1443. doi: 10.1001/jama.2009.1409.

Dakin HA, Devlin NJ, and Odeyemi IAO. (2006) "Yes", "No" or "Yes, but"? Multinomial modelling of NICE decision-making. *Health Policy (Amsterdam, Netherlands)* **77**(3), 352–367. doi: 10.1016/j.healthpol.2005.08.008.

Dakin H, Devlin N, Feng Y, Rice N, O'Neill P, and Parkin D. (2015) The influence of cost-effectiveness and other factors on NICE decisions. *Health Economics* **24**(10), 1256–1271. doi: 10.1002/hec.3086.

Degrassat-Théas A, Paubel P, Curzon OPd, Le Pen C, and Sinègre M. (2013) Temporary authorization for use: does the French patient access programme for unlicensed medicines impact market access after formal licensing?' *PharmacoEconomics* **31**(4), 335–343. doi: 10.1007/s40273-013-0039-4.

Devlin N, and Parkin D. (2004) Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. *Health Economics* **13**(5), 437–452. doi: 10.1002/hec.864.

Fischer KE, Rogowski WH, Leidl R, and Stollenwerk B. (2013) Transparency vs. closed-door policy: do process characteristics have an impact on the outcomes of coverage decisions? A statistical analysis. *Health Policy* **112**(3), 187–196. doi: 10.1016/j.healthpol.2013.04.011.

Fischer KE, Heisser T, and Stargardt T. (2016) Health benefit assessment of pharmaceuticals: an international comparison of decisions from Germany, England, Scotland and Australia. *Health Policy* **120**(10), 1115–1122. doi: 10.1016/j.healthpol.2016.08.001.

Grepstad M, and Kanavos P. (2015) A comparative analysis of coverage decisions for outpatient pharmaceuticals: evidence from Denmark, Norway and Sweden. *Health Policy (Amsterdam, Netherlands)* **119**(2), 203–211. doi: 10.1016/j.healthpol.2014.12.013.

Harris AH, Hill SR, Chin G, Jing Li J, and Walkom E. (2008) The role of value for money in public insurance coverage decisions for drugs in Australia: a retrospective analysis 1994-2004. *Medical Decision Making: An International Journal of the Society for Medical Decision Making* **28**(5), 713–722. doi: 10.1177/0272989X08315247.

Kanavos P, Angelis A, and Drummond M. (2019) An EU-wide approach to HTA: an irrelevant development or an opportunity not to be missed?' *The European Journal of Health Economics* **20**(3), 329–332. doi: 10.1007/s10198-019-01037-2.

Kanavos P, Visintin E, and Gentilini A. (2023) Algorithms and heuristics of health technology assessments: a retrospective analysis of factors associated with HTA outcomes for new drugs across seven OECD countries. *Social Science & Medicine* **331**(August), 116045. doi: 10.1016/j.socscimed.2023.116045.

Kristensen FB. (2012) Development of European HTA: from vision to EUnetHTA. *Michael* **9**, 147–156.

Le Pen C, Priol G, and Lilliu H. (2003) What criteria for pharmaceuticals reimbursement?' *The European Journal of Health Economics* **4**(1), 30–36. doi: 10.1007/s10198-002-0145-2.

Maguire B, and Tilden D. (2013) Key drivers of Pbac decisions for the reimbursement of orphan drugs on the life saving drugs Program. *Value in Health* **16**(7). Elsevier: A480–A481. doi: 10.1016/j.jval.2013.08.916.

Main C, Schäfer C and Kanavos P (2025) From vision to reality: the eu's pharmaceutical reforms and the path to improved access. *PharmacoEconomics Open* **9**, 331–339. https://doi.org/10.1007/s41669-024-00556-w

Martelli F, Torre GL, Ghionno ED, Staniscia T, Neroni M, Cicchetti A, Bremen K V, Ricciardi W, and **NI-HTA Collaborative Group**. (2007) Health technology assessment agencies: an international overview of organizational aspects. *International Journal of Technology Assessment in Health Care* **23**(4), 414–424. doi: 10.1017/S026646230707064X.

Mathes T, Jacobs E, Morfeld J-C, and Pieper D. (2013) Methods of international health technology assessment agencies for economic evaluations- a comparative analysis. *BMC Health Services Research* **13**(1), 371. doi: 10.1186/1472-6963-13-371.

Maynou L, and Cairns J. (2019) What is driving HTA decision-making? Evidence from cancer drug reimbursement decisions from 6 European countries. *Health Policy* **123**(2), 130–139. doi: 10.1016/j.healthpol.2018.11.003.

McMillan HJ, and Campbell C. (2017) We need a "made in Canada" orphan drug framework. *CMAJ : Canadian Medical Association Journal* **189**(41), E1274–E1275. doi: 10.1503/cmaj.170195.

Nicod E. (2017) 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. *The European Journal of Health Economics: HEPAC: Health Economics in Prevention and Care* **18**(6), 715–730. doi: 10.1007/s10198-016-0823-0.

Nicod E, and Kanavos P. (2012) Commonalities and differences in HTA outcomes: a comparative analysis of five countries and implications for coverage decisions. *Health Policy (Amsterdam, Netherlands)* **108**(2–3), 167–177. doi: 10.1016/j.healthpol.2012.09.012.

Nicod E, and Kanavos P. (2016) Developing an evidence-based methodological framework to systematically compare HTA coverage decisions: a mixed methods study. *Health Policy (Amsterdam, Netherlands)* **120**(1), 35–45. doi: 10.1016/j.healthpol.2015.11.007.

Nicod E, Maynou L, Visintin E, and Cairns J. (2020) Why do health technology assessment drug reimbursement recommendations differ between countries? A parallel convergent mixed methods study. *Health Economics, Policy and Law* **15**(3). Cambridge University Press: 386–402. doi: 10.1017/S1744133119000239.

Norton EC, Dowd BE, Maciejewski ML. (2019) Marginal effects—quantifying the effect of changes in risk factors in logistic regression models. *JAMA* **321**(13), 1304–1305. doi: 10.1001/jama.2019.1954

OECD. (2020) Addressing Challenges in Access to Oncology Medicines'. Analytical Report. https://www.oecd.org/health/health-systems/Addressing-Challenges-in-Access-to-Oncology-Medicines-Analytical-Report.pdf.

O'Mahony JF (2019) Beneluxa: what are the prospects for collective bargaining on pharmaceutical prices given diverse health technology assessment processes?' *PharmacoEconomics* **37**(5), 627–630. doi: 10.1007/s40273-019-00781-w.

Rawlins MD, The changing nature of regulation: Its impact on health technology assessment. Unpublished monograph based on a Keynote Address at the London School of Economics, 2016.

Salas-Vega S, Bertling A, and Mossialos E. (2016) A comparative study of drug listing recommendations and the decision-making process in Australia, the Netherlands, Sweden, and the UK. *Health Policy* **120**(10), 1104–1114. doi: 10.1016/j.healthpol.2016.08.006.

Segal L, Dalziel K, and Mortimer D. (2010) Fixing the game: are between-silo differences in funding arrangements handicapping some interventions and giving others a head-start?' *Health Economics* **19**(4), 449–465. doi: 10.1002/hec.1483.

Srikanthan A, Penner N, Chan KKW, Sabharwal M, and Grill A. (2018) Understanding the reasons for provincial discordance in cancer drug funding—a survey of policymakers. *Current Oncology* **25**(4). Multidisciplinary Digital Publishing Institute: 257–261. doi: 10.3747/co.25.3993.

'StataCorp. 2019. Stata Statistical Software: Release 16. College Station, TX: StataCorp LLC.' n.d.

Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, Cates CJ, et al. (2019) RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* **366**, 1–8. doi: 10.1136/bmj.l4898.

Vreman RA, Mantel-Teeuwisse AK, Hövels M, Leufkens HGM, and Goettsch WG. (2020) Differences in health technology assessment recommendations among European jurisdictions: the role of practice variations. *Value in Health* **23**(1), 10–16. doi: 10.1016/j.jval.2019.07.017.

Wood EM, and Hughes DA. (2020) The new and non-transparent cancer drugs fund. *PharmacoEconomics* **38**(1), 1–4. doi: 10.1007/s40273-019-00871-9.