Title: Association between the use of surrogate measures in pivotal trials and health technology assessment decisions: A retrospective analysis of NICE and CADTH reviews of cancer drugs

Running title: Surrogate measures in HTA

Authors:

Ashlyn Pinto, MSc

Department of Health Policy

London School of Economics and Political Science

London

United Kingdom

Huseyin Naci, MHS PhD

Department of Health Policy

London School of Economics and Political Science

London

United Kingdom

Emilie Neez, MSc

Department of Health Policy

London School of Economics and Political Science

London

United Kingdom

Professor Elias Mossialos, MD PhD

Department of Health Policy

London School of Economics and Political Science

London

United Kingdom

Full address for correspondence:

Dr. Huseyin Naci, LSE Health, Cowdray House, London School of Economics and Political Science, 20 Houghton Street, London, WC2A 2AE, United Kingdom, Tel: +44 (0)20 7955 6874, E-mail: h.naci@lse.ac.uk

Funding/support:

This study did not receive funding.

Financial Disclosure:

None reported.

Precis: Despite criticism pointed at regulators for their increasing reliance on surrogate measures when approving new cancer medicines, decisions of HTA bodies are not negatively influenced by the use of surrogate measures.

Word count: 3815

Number of pages: 103 (84 from Appendices)

Number of tables: 2

Number of figures: 3

Appendices:

Pages: 84

Tables: 7

Figures: 1

Abstract

Objective: To assess whether using surrogate versus patient-relevant endpoints in pivotal trials of cancer drugs was associated with HTA recommendations in England (National Institute for Health and Care Excellence - NICE) and Canada (Canadian Agency for Drugs and Technologies in Health - CADTH).

Methods: Cancer drug approvals from 2012 to 2016 were categorised by demonstrating benefit on overall survival (OS), progression-free survival, disease response, or having no comparator. Approvals were analysed by benefit category and HTA recommendation. The association between benefit (surrogate versus OS) and recommending a drug was examined using descriptive statistics and linear probability models controlling for unmet need, orphan designation, and cost-effectiveness.

Results: Of 42 cancer indications that NICE recommended, 15 (36%) demonstrated OS benefit. Cancer indications with OS benefit were less likely to receive a recommendation from NICE than those without (*p=0.04*). In linear probability models, availability of OS benefit was no longer associated with a recommendation from NICE (*p=0.32*). Cost-effective cancer drugs had a 55.6% [*95% CI*: 38.9% to 72.3%] higher probability of receiving a recommendation from NICE than those that were not. In Canada, 15 of 37 (41%) cancer indications that were recommended showed OS benefit. There was no detectable association between surrogate measures and CADTH recommendations based on descriptive statistics (*p=0.62*) or in linear probability models (*p=0.73*).

Conclusions: When cost-effectiveness was considered, pivotal trial endpoints were not associated with NICE recommendations. Pivotal trial endpoints, unmet need, orphan status, and cost-effectiveness did not explain CADTH recommendations.

Key Words: surrogate outcomes, health technology assessment, cancer medicines

Highlights

- Drug licensing agencies are under increasing pressure to expedite the approval of new cancer medicines. A common strategy to facilitate faster development of therapeutic agents is to use surrogate measures of clinical benefit. The validity of surrogates of overall survival in many cancer indications has been debated. There is no up to date study that evaluates the impact of surrogate measures on HTA recommendations.
- This study shows no evidence that the use of surrogate measures in cancer trials are associated with a HTA decision to reject a drug. During our study period, cancer drugs without overall survival benefit were more likely to be recommended by NICE in England. However, when taking into consideration multiple factors, overall survival benefit did not affect NICE's decisions, while cost-effectiveness did. CADTH decisions in Canada were not associated with the use of surrogate endpoints, unmet need, cost-effectiveness or orphan status.
- These findings raise questions about the assumptions made in cost-effectiveness models in England, as multiple systematic reviews in oncology have found no clear relationship between surrogate measures and patient-centered outcomes, namely overall survival and quality-of-life.
- The observed variation in CADTH decisions is not explained by known and measured factors in our analysis. Further clarity is needed on the framework guiding CADTH decisions on new cancer drugs.

Introduction

Surrogate measures, such as progression-free survival (PFS), are substitutes to patient-relevant endpoints in cancer trials.[1] Thus, changes seen in a surrogate endpoint are expected to emulate changes in a patient-relevant endpoint, such as overall survival (OS) or quality of life (QoL). Using surrogate measures in cancer trials have several benefits as they allow for shorter study durations, smaller sample sizes,[2] lower costs and earlier approval of new drugs for cancer patients.[3,4] In recent years, PFS and response rates have become increasingly popular surrogate measures used in cancer trials.[5]

However, evidence shows that such surrogate measures may be an unpredictable marker of OS[6] and QoL benefit.[7] A systematic review in 2018[5] failed to find a significant association between PFS and QoL in cancer trials. Another review of trial-level meta-analyses in 2019[8] showed a low to moderate correlation between surrogate measures and OS.

Licensing agencies such as the European Medicines Agency (EMA) focus on a drug's benefit-risk profile for granting marketing authorisations.[1,4] However, Health Technology Assessment (HTA) bodies, such as the National Institute of Health and Care Excellence (NICE) in England and the Canadian Agency for Drugs and Technologies in Health (CADTH) in Canada, are concerned with long-term comparative clinical and cost-effectiveness in addition to efficacy and safety. HTA decision making is a multidisciplinary process that includes many factors that may differ across individual agencies. While some contributing factors are difficult to ascertain, such as the extent of input from experts and patients, agencies have a set of explicit criteria to inform their decisions. For example, both CADTH[9] and NICE[10] consider unmet need, disease prevalence, and equity alongside data on the drug's comparative efficacy, safety, and cost-effectiveness.

In addition, endpoints used in pivotal trials may influence HTA decisions. Indeed, HTA bodies have a clear preference for patient-relevant endpoints over surrogates.[1,11,12] According to a study[13] of 21 international HTA agencies, most organisations' guidelines stated that a surrogate measure was acceptable only in exceptional situations, given its validity had been proven based on factors such as consistently strong statistical association and biological and pathophysiological plausibility. This was further reiterated in a study[14] highlighting that HTA agencies preferred systematic reviews to confirm the association between a surrogate and patient-relevant endpoint when making decisions about reimbursement. Despite these preferences, a study[15] exploring the use of surrogate measures in cost-effectiveness models within HTA agencies found that submitted reports often lacked this evidence, yielding uncertainty surrounding the surrogate measures.

The objective of this study was to assess whether the use of surrogate measures in pivotal trials of cancer drugs authorised by the EMA affected funding recommendations made by HTA bodies in England (NICE) and Canada (CADTH). These two countries share transparency in their appraisal processes as well as publically funded healthcare systems.

Methods

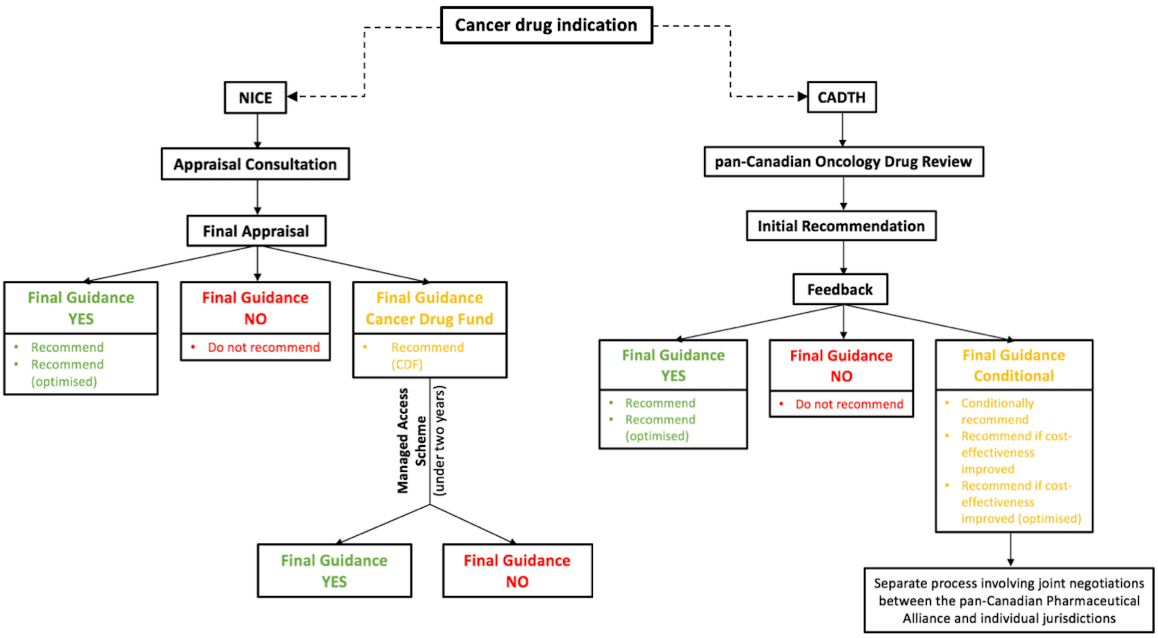
Two independent investigators used the EMA's website to retrieve all marketing authorisations for human cancer drugs approved between 1st January 2012 and 31st December 2016. All drugs under 'Cancer' and its subcategories, such as 'Neoplasms by site', were explored. Drugs were eligible for review if they were indicated specifically for cancer treatment. Generic drugs, hybrid medicines, extensions for indications and other cancer types, and supportive cancer drugs were excluded. All European Public Assessment Reports (EPARs) for the included indications were systematically reviewed to identify and analyse pivotal studies. It was possible for one drug's market authorisation to have more than one indication. A Microsoft Excel data extraction form was created to record study characteristics from all pivotal studies, including trial design, sample size, availability of control group, blinding and crossover procedures, participant characteristics, duration of follow-up, primary and secondary endpoints, and trial results.

Two investigators then independently identified instances where these indications were considered in NICE and CADTH HTA reviews via their websites, up until November 2018. **Figure 1** displays policy pathways and potential recommendations for cancer drug indications in England and Canada. A separate Microsoft Excel data extraction form was created to systematically review HTA documents for each cancer drug, which included the indication(s), document number, and recommendation. In instances where appraisals were resubmitted, the initial guidance was only used if the resubmitted appraisal was still in progress.

Each HTA document was also assessed to see if the therapeutic indication fulfilled an unmet clinical need, was considered cost-effective, or had an orphan designation. These variables were the same as those considered in Dakin et. al's study[16] investigating the influence of multiple factors on NICE's decision-making. Unmet need was determined by agencies, often with input from patient advocacy groups, and was reported under the "Current Practice" and "Limitations of Current Therapy" in NICE and CADTH appraisals, respectively. Cost-effectiveness was discussed under the "Committee Discussions" section of appraisals for NICE, with a threshold of £30,000 per quality-adjusted life year (QALY)[17], and the "Economic Evaluation" section for CADTH, with no explicit threshold.[18,19] Budget impact analyses were not included in the analysis as

it is not an explicit criterion for NICE, and CADTH considers it alongside cost-effectiveness analyses. A drug was classified as "orphan" if it targeted a disease with a prevalence 5 or less per 10,000 in the European Union.

Figure 1: Policy pathway and potential recommendations for cancer drug indications in England and Canada



Two investigators then individually reviewed results of the pivotal trials (both primary and secondary endpoints) and categorised them into one of the following four groups: trials that demonstrated statistical significance in terms of (1) OS, (2) PFS, (3) disease response (DR) (e.g., response rate), and (4) those that showed no statistical evidence of benefit in comparative trials (i.e., trials with no comparator groups). Results were considered statistically significant if they met criteria set out in the 'Statistical methods' section of the pivotal study – typically a two-sided p-value with an alpha of 0.05. Consistent with previous studies[20], single-arm trials were considered to show no statistical significance as they did not compare the intervention against a placebo or alternative treatment. When an indication showed benefit in more than one category, it was hierarchically classified into the category of highest importance. OS benefit was considered the most persuasive outcome based on recent EMA guidelines on the evaluation of anticancer drugs,[21] followed by PFS, DR, and lastly no statistical benefit.

Drug indications were then analysed according to both benefit categorisation and HTA recommendation from NICE and/or CADTH. Paired t-tests were used to compare the likelihood of HTA agencies recommending a drug between cancer indications with and without OS benefit at the time of EMA approval. Linear probability models were then used to compare the likelihood of recommending a drug based on OS benefit, unmet need, orphan designations, and cost-effectiveness. Lastly, disagreements in recommendations between NICE and CADTH for the same indication were narratively analysed.

Results

Cancer medicines approved by the EMA between 2012 and 2016

The EMA approved 58 indications for 48 drugs between 2012 and 2016. Approved indications and respective pivotal study data can be found in the online-only supplementary table (S1). All pivotal studies had OS and DR related endpoints, primarily objective response-rate, as either a primary or secondary outcome. All pivotal trials, except for four, reported PFS as either a primary or secondary outcome. Of 58 indications, 25 (43%) demonstrated statistical evidence of OS benefit, 35 (60%) PFS benefit, and 39 (67%) DR benefit. 17 (29%) of the 58 indications' pivotal studies were single-arm trials. Based on the hierarchal classification of benefit, 25 (43%) of the 58 indications were categorised as exhibiting OS benefit, 14 (24%) PFS benefit, 2 (3%) DR benefit, and 17 (29%) no comparator. **Table 1** displays EMA approved indications and their respective HTA recommendations.

Table 1: EMA cancer medicines market authorisations from 2012 to 2016, categorised by benefit shown when authorised and NICE and CADTH guidance

ACTIVE SUBSTANCE	BENEFIT CATEGORY	NICE GUIDANCE	CADTH GUIDANCE
Lenvatinib (2015): DTC	OS	Recommended (Optimised)	Recommend if cost-effectiveness improved
Necitumumab (2016): nSCLC	OS	Not recommended	-
Nintedanib (2014): nSCLC	OS	Recommended	-
irinotecan hydrochloride tihydrate (2016): PMA	OS	Not recommended	Recommend if cost-effectiveness improved
Ramucirumab (2014): combination; gastric	OS	Not recommended	Recommend if cost-effectiveness improved
Ramucirumab (2014): monotherapy; gastric	OS	Not recommended	Not Recommended
trifluridine/tipiracil (2016): CRC	OS	Recommended	Not Recommended
Regorafenib (2013): CRC	OS	Terminated appraisal	Not recommended

ACTIVE SUBSTANCE	BENEFIT CATEGORY	NICE GUIDANCE	CADTH GUIDANCE
trastuzumab emtansine (2013): BC	OS	**Recommended**	**Recommend if cost-effectiveness improved**
Pertuzumab (2013): BC	OS	**Recommended**	**Recommend if cost-effectiveness improved**
Radium Ra223 dichloride (2013): PC	OS	**Recommended (Optimised)**	**-**
Enzalutamide (2013): PC	OS	**Recommended**	**Recommended**
Cobimetinib (2015): melanoma	OS	**Not recommended**	**Recommend (Optimised) if cost-effectiveness improved**
Nivolumab (2015): melanoma	OS	**Recommended**	**Recommend (Optimised) if cost-effectiveness improved**
Vemurafenib (2012): melanoma	OS	**Recommended**	**Recommended (Optimised) if cost-effectiveness improved**
Olaratumab (2016): soft tissue sarcoma	OS	**Recommended (CDF)**	**Conditionally recommended if cost-effectiveness improved**
Decitabine (2012): AML	OS	**Terminated Appraisal**	**-**
Pegaspargase (2016): ALL	OS	**Recommended (Optimised)**	**-**
Ibrutinib (2014): CLL	OS	**Recommended**	**Recommend if cost-effectiveness improved**
Idelalisib (2014): CLL; one prior therapy	OS	**Recommended (Optimised)**	**Recommend if cost-effectiveness improved**
Obinutuzumab (2014): CLL	OS	**Recommended (Optimised)**	**Recommended**
Elotuzumab (2016): MM	OS	**In progress**	**-**
Carfilzomib (2015): MM	OS	**Discontinued**	**Recommend (Optimised) if cost-effectiveness improved**
Pomalidomide (2013): MM	OS	**Recommended (Optimised)**	**Recommend if cost-effectiveness improved**
Aflibercept (2013): CRC	OS	**Not Recommended**	**Not Recommended**
Vandetanib (2012): MTC	PFS	**In progress**	**Recommend if cost-effectiveness improved**
Afatinib (2013): nSCLC	PFS	**Recommended (Optimised)**	**Recommended (Optimised)**
Olaparib (2014): EOFP	PFS	**Recommended (Optimised)**	**Recommend if cost-effectiveness improved**
Palbociclib (2016): BC; combination with AI	PFS	**Recommended**	**Recommend (Optimised) if cost-effectiveness improved**
Palbociclib (2016): BC; combination with fulvest	PFS	**Suspended**	**In progress**
Axitinib (2012): RCC	PFS	**Recommended**	**Recommended**
Pembrolizumab (2015): melanoma	PFS	**Recommended (Optimised)**	**Recommend (Optimised) if cost-effectiveness improved**
Trametinib (2014): melanoma	PFS	**Recommended (Optimised)**	**Recommend if cost-effectiveness improved**
Dabrafenib (2013): melanoma	PFS	**Recommended**	**Recommend if cost-effectiveness improved**
Ixazomib (2016): MM	PFS	**Recommended (Optimised and CDF)**	**In progress**
Panobinostat (2015): MM	PFS	**Recommended**	**-**
Pixantrone (2012): nHBL	PFS	**Recommended (Optimised)**	**-**
Cabozantinib (2014): MTC	PFS	**Recommended (CDF)**	**-**
Idelalisib (2014): CLL; first-line	PFS	**Recommended**	**Recommend if cost-effectiveness improved**

ACTIVE SUBSTANCE	BENEFIT CATEGORY	NICE GUIDANCE	CADTH GUIDANCE
Talimogene laherparepvec (2015): melanoma	DR	Recommended (Optimised)	-
Bosutinib (2013): CML	DR	Recommended	Recommend if cost-effectiveness improved
Osimertinib (2016): nSCLC	Single-arm Trial	Recommended (Optimised and CDF)	Recommend if cost-effectiveness improved
Ceritinib (2015): nSCLC	Single-arm Trial	Recommended	Recommend if cost-effectiveness improved
Crizotinib (2012): nSCLC	Single-arm Trial	Recommended	Recommend if cost-effectiveness improved
Sonidegib (2015): BCC	Single-arm Trial	-	-
Vismodegib (2013): BCC; metastatic	Single-arm Trial	Not Recommended	Recommend if cost-effectiveness improved
Vismodegib (2013): BCC; locally advanced	Single-arm Trial	Not Recommended	Recommend if cost-effectiveness improved
Ponatinib (2013): CML	Single-arm Trial	Recommended	Recommend if cost-effectiveness improved
Ponatinib (2013): ALL	Single-arm Trial	Recommended	Recommend if cost-effectiveness improved
Blinatumomab (2015): ALL	Single-arm Trial	Recommended	Recommend if cost-effectiveness improved
Ibrutinib (2014): MCL	Single-arm Trial	Recommended (Optimised)	Recommend if cost-effectiveness improved
Idelalisib (2014): FL	Single-arm Trial	In progress	Not recommended
Daratumumab (2016): MM	Single-arm Trial	Recommended (Optimised and CDF)	Not recommended
Brentuximab vedotin (2012): HL; post-ASCT	Single-arm Trial	Recommended	Recommend if cost-effectiveness improved
Brentuximab vedotin (2012): HL; non-ASCT candidates	Single-arm Trial	Recommended	Not recommended
Brentuximab vedotin (2012): sALCL	Single-arm Trial	Recommended (Optimised)	Recommend if cost-effectiveness improved
Venetoclax (2016): CLL; no 17p deletion	Single-arm Trial	Recommended (CDF)	Conditionally recommended if cost-effectiveness improved
Venetoclax (2016): CLL; with 17p deletion	Single-arm Trial	Recommended (CDF)	Not Recommended

AI = Aromatase Inhibitor, ALL = Acute Lymphoblastic Leukaemia , AML = Acute Myeloid Leukaemia, ASCT = Autologous Stem Cell Transplant, BC = Breast Cancer, BCC = Basal Cell Carcinoma, CDF = Cancer Drug Fund, CLL = Chronic Lymphocytic Leukaemia, CML = Chronic Myeloid Leukaemia, CRC = Colorectal Cancer, DTC = Differentiated Thyroid Carcinoma, EOFP = Epithelial Ovarian, Fallopian Tube, or Primary Peritoneal Cancer, FL = Follicular Lymphoma, HL = Hodgkin Lymphoma, MCL = Mantle Cell Lymphoma, MM = Multiple Myeloma, MTC = Medullary Thyroid Cancer, nHBL = Non-Hodgkin B-Cell Lymphoma, nSCLC = Non-Small Cell Lung Cancer, PC = Prostate Cancer, PMA = Pancreatic Metastatic Adenocarcinoma, RCC = Renal Cell Carcinoma, sALCL = Systemic Anaplastic Large Cell Lymphoma

NICE and CADTH recommendations

Of 58 EMA-approved indications, 57 were subsequently reviewed by NICE: it recommended 45 (79%) indications - 17 of which had an "Optimised" recommendation and 7 under the Cancer Drugs Fund (CDF), did not recommend 8 (14%), was reviewing 3 (5%), terminated 2 appraisals, suspended 1, and discontinued 1 (7%). Specific NICE guidance for each indication can be found in the online-only supplementary table (S2).

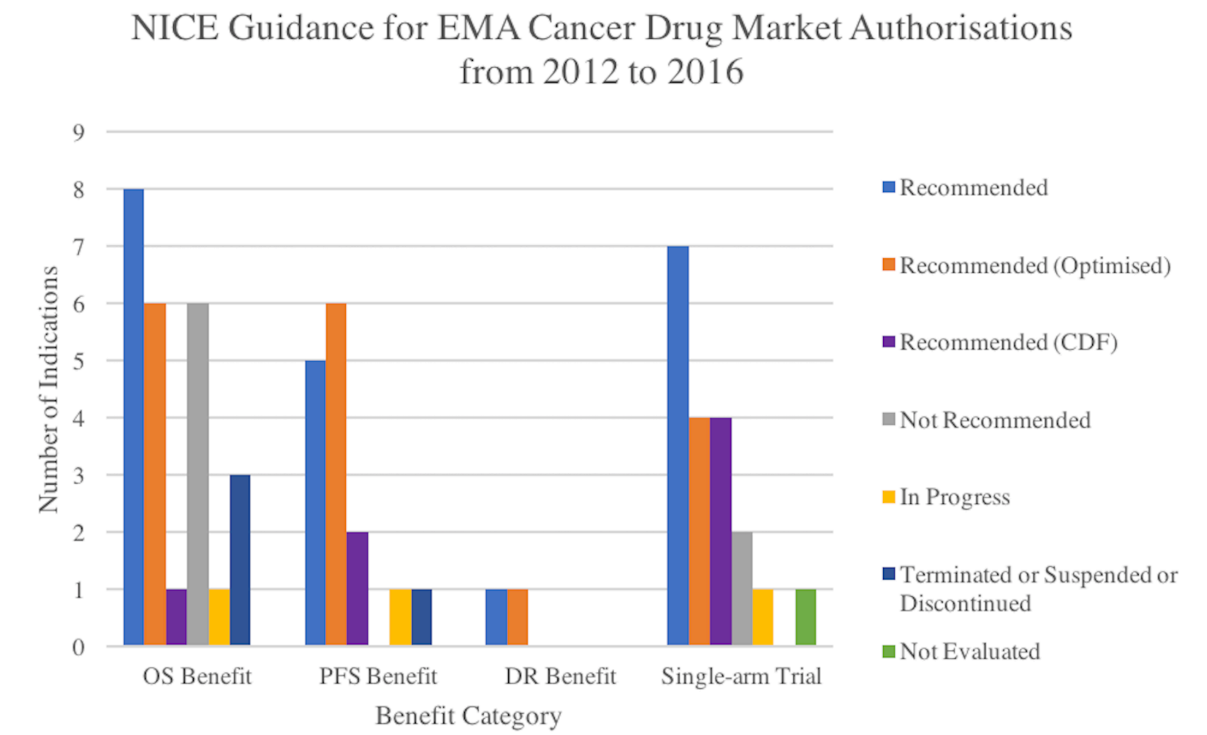
CADTH provided guidance on 47 of the 58 indications approved by the EMA: it recommended 37 (79%) - 1 of which was an "Optimised" recommendation, 25 of which were recommended if the cost-effectiveness was improved, 6 that were "Optimised" and if cost-effectiveness was improved, and 2 of which were conditionally recommended. 8 (17%) indications were not recommended and 2 (4%) were in progress. Specific CADTH guidance for each indication can be found in the online-only supplementary table (S3), as well as a cross-tab of the number of indications by recommendation for both agencies (S4). Detailed reasons for not recommending indications for each agency can be found in supplementary tables 5 and 6.

NICE and CADTH recommendations per benefit category

Of the 25 indications based on OS benefit, 8 (32%) were recommended, followed by 6 (24%) recommended (optimised) and 6 (24%) not recommended by NICE; of the 14 based on PFS data 6 (43%) were recommended (optimised), followed by 5 (36%) recommended; the 2 based on DR data were recommended; of the 17 with no statistical benefit, 7 (41%) were recommended followed by 4 (24%) recommended (optimised) and 4 (24%) recommended (CDF) (**Figure 2**).

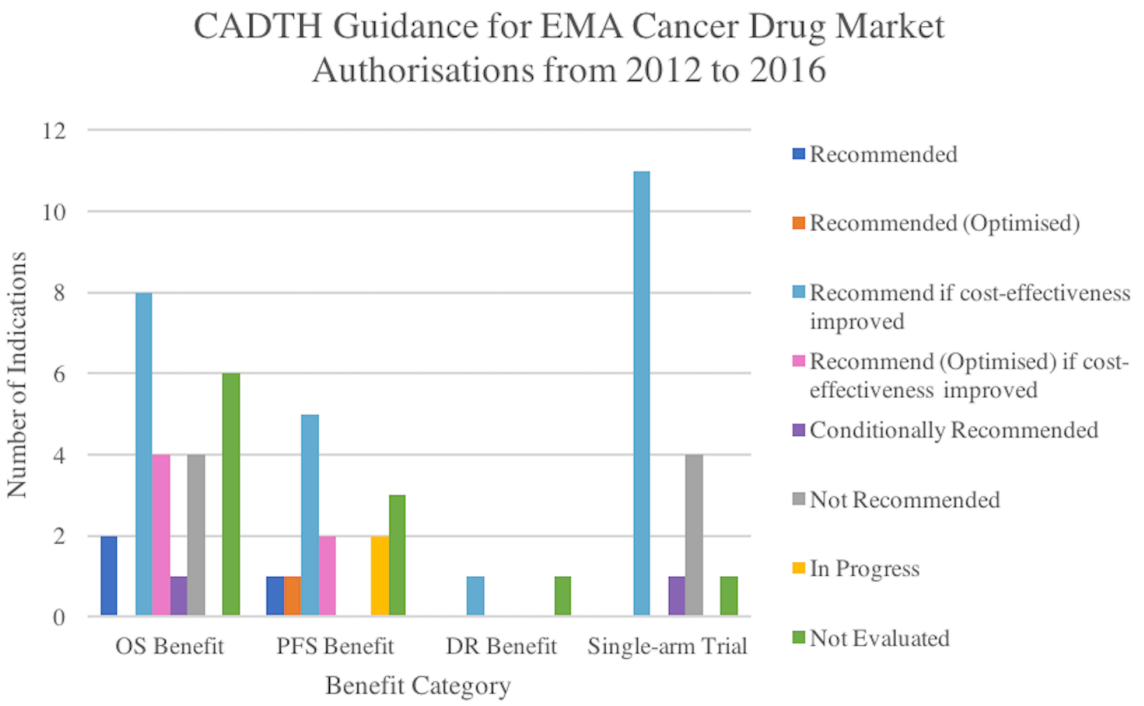
Excluding drugs that were in progress, terminated, suspended, discontinued or not evaluated, NICE provided guidance for 50 indications. NICE recommended 15 of 21 (71%) indications that had OS benefit at the time of EMA approval, and 27 of 29 (93%) indications that were approved on the basis of surrogate measures (PFS, DR, and single-arm trial data). The difference in proportions was 22 percentage points (71% versus 93%, *p=0.04*). When unmet need, orphan designation, and cost-effectiveness were incorporated into a linear probability model, OS benefit was no longer a statistically significant predictor (p=0.32). Unmet need (p=0.38) and orphan designation (p=0.09) were also not statistically significant predictors. Cost-effectiveness was the only significant predictor (p=0.00), increasing a drug's probability of NICE recommending it by 55.6% [95% CI = 38.9 to 72.3%].

Figure 2: NICE guidance for EMA cancer drug market authorisations from 2012 to 2016



Of the 25 indications based on OS data for CADTH, 8 (32%) were recommended if the cost-effectiveness was improved to an acceptable level, followed by 6 (24%) not evaluated; of the 14 based on PFS data, 5 (36%) were recommended contingent on cost-effectiveness being improved, followed by 3 (21%) not evaluated; of the two based on DR data, 1 (50%) was not evaluated and 1 (50%) was recommended if cost-effectiveness improved; of the 17 with no statistical benefit, 11 (65%) were recommended if cost-effectiveness was improved, followed by 4 (24%) that were not recommended (**Figure 3**). Online-only supplementary data show the proportion of specific recommendations per benefit category for NICE and CADTH (S7).

Figure 3: CADTH guidance for EMA cancer drug market authorisations from 2012 to 2016



Excluding drugs that were in progress or not evaluated, CADTH provided guidance for 45 indications. CADTH recommended 15 of 19 (79%) indications based on OS benefit, and 22 of 26 (85%) based on non-OS benefit. The difference in proportions was 6 percentage points (85% versus 79%, *p=0.62*). When unmet need, orphan designation, and cost-effectiveness were incorporated into a linear probability model, availability of OS was still not a statistically significant predictor (p=0.73). Unmet need (p=0.62), orphan designation (p=0.72) and cost-effectiveness (p=0.65) were also not statistically significant predictors for a drug being recommended by CADTH. **Table 2** provides the complete results of the linear probability model for NICE and CADTH.

Table 2: NICE and CADTH Linear Probability Model for EMA Cancer Drug Market Authorisations from 2012 to 2016

NICE and CADTH Linear Probability Model for Recommending EMA Cancer Drug Market Authorisations from 2012 to 2016			
NICE*			
	Probability	**P-value**	**95% Confidence Interval**
Overall Survival	-7.91%	0.318	[-23.67, 7.85]
Unmet Need	8.98%	0.381	[-11.45, 29.41]
Orphan Designation	13.68%	0.086	[-2.01, 29.38]
Cost-effectiveness	55.64%	0.000	[38.95, 72.33]
CADTH**			
	Probability	**P-value**	**95% Confidence Interval**
Overall Survival	-4.31%	0.733	[-29.69, 21.07]
Unmet Need	-13.48%	0.618	[-67.59, 40.63]
Orphan Designation	-4.70%	0.723	[-31.33, 21.92]
Cost-effectiveness	13.56%	0.653	[-46.97, 74.09]

***p-value associated with F-value (Probability > F) = 0.000. R^2 = 56%.**
****p-value associated with F-value (Probability > F) = 0.8530. R^2 = 3.2%.**

Out of the 58 indications, 47 (81%) were evaluated by both NICE and CADTH. The HTA bodies disagreed on 9 (19%) indications, specifically with one agency recommending the drug while the other did not. For 5 of 9 (56%) indications that NICE did not recommend, NICE concluded that the incremental cost-effectiveness ratio (ICER) was over the threshold, while CADTH recommended the drugs only if the cost-effectiveness was improved. For 3 of 9 (33%) indications that NICE did not recommend, NICE had uncertainty about clinical benefit, whereas CADTH was satisfied with the evidence. When CADTH did not recommend a drug, the main reason for 3 of 9 (33%) indications was that CADTH had clinical and cost-effectiveness uncertainty, while NICE recommended it under the CDF to collect more data. 2 of 9 (22%) indications not recommended by CADTH were not cost-effective, while for NICE these drugs reached End-of-Life criteria and had a higher cost threshold. Supplementary table S8 lists the reasons for each HTA disagreement.

Discussion

This study identified 58 cancer indications that received EMA marketing authorisations between 2012 and 2016. Of these, less than half were approved based on statistically significant OS benefit. NICE issued guidance on 50 of the 58 market authorisations. Of the 42 recommended drugs, less than half were based on OS benefit. During our study period, NICE was more likely to recommend drugs with non-OS benefit than OS benefit. However, when considering OS in conjunction with unmet need, orphan designation and cost-effectiveness, there was no difference between the probability of drugs with and without OS benefit to be recommended. Instead, cost-effectiveness was a statistically significant factor, consistent with Dakin et. al's[16] findings that cost-effectiveness was the most important predictor for a drug being recommended. CADTH issued guidance on 45 of the 58 EMA market authorisations. Only 3 indications were solely recommended, while 34 were recommended with stipulations such as cost-effectiveness being improved or optimising the original indication. On its own or when analysed with unmet need, orphan designation, and cost-effectiveness, availability of OS benefit did not effect the likelihood of CADTH recommending a drug.

In absolute terms, NICE was more likely to recommend drugs based on surrogate benefit, specifically PFS, rather than OS. An in-depth review showed the main reasons for this being that crossover diminished OS benefit, studies were not powered for OS, results suggested improvement despite non-significance, and results of ongoing trials were still pending. However, from the results of the linear probability model, the single most important factor predicting a recommendation from NICE was cost-effectiveness, as the difference between drugs with OS benefit and those without became no longer statistically significant. These results raise questions about the assumptions made in cost-effectiveness models when extrapolating the findings of surrogate measures to long-term clinical improvements; many systematic reviews[5-7] found no clear relationship between surrogate endpoints and OS or QoL.

OS benefit did not increase the likelihood of CADTH recommending a drug, which is inconsistent with the literature[2,11,22] that HTA bodies prefer endpoints that show direct clinical improvements. A study by Kreeftmeijer et al.[23] found that patient-relevant endpoints in cancer clinical trials were associated with a higher percentage of being recommended by European payers, specifically in France, Germany and the United Kingdom. However, the authors also found that statistical significance alone was insufficient, as reimbursement bodies also focused on clinically relevant differences.[23] This aligns with our study results, as OS was not a significant predictor for a CADTH recommending a drug by itself or when accounting for unmet need, cost-effectiveness or orphan designation. Previous literature[24] found that differences in drug recommendations

between England and Canada may be less about clinical or economic evidence, but instead associated with agency-specific processes such as willingness to accept risk. More clarity on the framework guiding CADTH decisions on cancer drugs is warranted.

Eight indications were not recommended by NICE and eight by CADTH. Interestingly, 6 of 8 (75%) of the medicines that were not recommended by NICE demonstrated OS benefit. A common reason for not recommending these drugs was high ICERs, some being over £100,000 per QALY gained, exceeding NICE's cost-effectiveness threshold of £30,000.[17] This is consistent with a previous review published in 2007[25], which found that most NICE rejections were for drugs exceeding this threshold. Our findings suggest that statistically significant OS benefit does not outweigh cost-effectiveness – an important factor to consider in a publicly funded healthcare system. Similarly, rejection recommendations by CADTH were all cost-ineffective. These findings for both agencies are supported by Lim et al.[26], who found that countries financed by general taxation are more likely to reimburse drugs the lower their ICER is, whereas countries such as Germany solely focus on clinical evidence. Although half of the rejection recommendations by CADTH showed statistical OS benefit, the magnitude of benefit was modest. Recommendation disagreements between NICE and CADTH commonly had one agency not recommending an indication while the other put a stipulation on its recommendation. For example, when CADTH did not recommend a drug due to clinical uncertainty, NICE would recommend it under the CDF. Moreover, when NICE did not recommend a drug because the ICER was above its threshold, CADTH would recommend it only if the cost-effectiveness was improved to an acceptable level.

Of the 58 cancer indications approved by the EMA, NICE evaluated 57 (98%), while CADTH only evaluated 47 (81%). Discrepancies in this can be explained by the fact that the EMA is a European regulatory agency, to which NICE directly looks for market authorisations. CADTH was used as a comparison HTA agency as the healthcare systems are similar, but has its own regulatory agency – the Health Products and Food Branch of Health Canada. Further, a major difference between NICE and CADTH was how each categorised drugs they recommended. While NICE either recommended or optimised a recommendation, CADTH, in addition to these, recommended a drug contingent on its cost-effectiveness being improved to an acceptable level. Of the 58 indications, 31 (53%) fell under this category for CADTH, while only 3 (5%) indications were solely recommended. As CADTH has not publicly stated its threshold[18,19], what they deem an acceptable level of cost-effectiveness is unknown. It is important to note that CADTH's 'Recommend if cost-effectiveness is improved' guidance could potentially be a 'Not recommended' guidance from NICE, as their approval processes slightly differ. CADTH has an initial and final recommendation. After the initial recommendation, all

participatory stakeholders may comment on the guidance. Once a final recommendation is issued, a separate process involves the pan-Canadian Pharmaceutical Alliance supporting individual jurisdictions through negotiations to ensure value for publicly funded drug programs.[27] Comparatively, NICE recommends that pharmaceutical companies offer patient access schemes at the time of submission. A review of guidance that rejects a drug may occur if there is a discount in price (manufacturers can engage in confidential negotiations), but this would result in a new appraisal that reviews the initial decision.[28] Further, differing appeal processes may account for variations between NICE[29] and CADTH.[30]

When analysing NICE's approval process it is important to consider the Cancer Drug Fund (CDF) – a source of funding for cancer drugs in England that provides faster access to new treatments. From its inception in 2011 to 2016, the CDF reviewed treatments not yet appraised or considered too rare for appraisal by NICE, not recommended by NICE, or not licensed for the proposed clinical purpose.[31] There has been criticism that NICE strategically did not recommend drugs knowing that the CDF would fund them instead. However, the CDF was restructured in 2016 to operate under NICE.[32] Presently, for the CDF to consider a drug, NICE must first recommend it for use within the CDF. This recommendation means the drug has potential to be eligible to be recommended by NICE, but has remaining clinical, and thus cost-effectiveness, uncertainty.[32] The drug then enters a managed access period to resolve any significant remaining clinical uncertainty.[32] This paper's analysis of NICE's recommendations was completed both before and after the CDF's restructuring. Thus, there is the possibility that NICE strategically did not recommend certain indications with the expectation that the CDF would for some products in our sample.

This study has limitations. Firstly, EMA-approved drugs were analysed so that comparisons could be made between England and Canada based on one list of drugs, despite Canada having its own regulatory body. Secondly, we relied exclusively on publicly available information. We were unable to consider the extent of expert or patient contribution to decision-making, as these were difficult to accurately quantify from public reports. Also, it remains a possibility that individual factors and their relative contribution to HTA decisions may have evolved over our study period. Fourthly, this study defined benefit as statistical significance in a comparative trial. Thus, single-arm trials were categorised as having no statistical evidence of benefit, even when they demonstrated improvement in terms of surrogate measures. Our focus on statistically significant findings may over-simplify the complex HTA decision-making process, which also considers aspects beyond the type of endpoint that showed statistical benefit. HTA agencies also consider factors such as standard of care available, therapeutic area, and political factors[16], for example the relationship between the CDF and NICE.

Furthermore, clinically meaningful benefit is becoming increasingly important in oncology, with organisations such as ESMO[33] and ASCO[34] developing frameworks to evaluate it. These were not assessed as they are currently not used by NICE and CADTH, but do provide important information on value in cancer care. Of note, drugs with OS benefit do score higher under the ESMO framework than those without.[33] Despite many factors influencing the decision-making process, surrogate measures still pose a substantial challenge to HTA bodies due to uncertainty surrounding their correlation to clinical outcomes.

While drug licensing agencies increasingly rely on surrogate measures to grant marketing authorisation, HTA bodies have traditionally preferred patient-relevant endpoints.[12,15] In practice, HTA bodies must make decisions on licensed products even when desirable evidence is unavailable. Our results show that statistically significant benefit from a patient-relevant endpoint does not always confirm a HTA recommendation. Further, surrogate measures are becoming increasingly context-specific[2] with diverse patient populations and the emergence of new technologies and mechanisms of actions.[35] This makes evidence of surrogate validity from retrospective trials harder to apply to different contexts. Current surrogate measures used in cancer research do not show consistently strong associations with OS. HTA agencies must ensure that surrogate measures amount to meaningful benefit. One solution would be to capitalise on conditional reimbursements[25,36], which stipulate that drugs need to prove clinical benefit within a certain time frame after reimbursement.

Conclusion

The use of surrogate measures in cancer trials was not associated with a HTA decision to reject a drug in England and Canada. NICE was less likely to recommend cancer drugs with OS benefit compared to those with PFS or DR. However, when also considering unmet need, orphan designation and cost-effectiveness, OS was not statistically significantly associated with a drug being recommended. Instead, this was driven by cost-effectiveness. This finding raises questions about the assumptions made in cost-effectiveness models in England, as multiple systematic reviews have demonstrated an unclear relationship between PFS and OS or QoL. Pivotal trial endpoints, unmet need, orphan status, and cost-effectiveness did not significantly affect CADTH recommendations during our study period. Further clarity is needed on the factors guiding recommendations of cancer drugs in Canada to ensure accountability.

References

1. Fleming TR, Powers JH. Biomarkers and surrogate endpoints in clinical trials. *Stat Med*. 2012;31(25):2973-2984.
2. Clarke JM, Wang X, Ready NE. Surrogate clinical endpoints to predict overall survival in non-small cell lung cancer trials—are we in a new era? *Transl Lung Cancer Res*. 2015;4(6):804-808.
3. Fischer A, Hernandez-Villafuerte K, Latimer N, Henshall C. Challenges and methodologies in using progression free survival as a surrogate for overall survival in oncology. *Int J Technol Assess Health Care*. 2018;34(3):300-316.
4. Rocchi A, Khoudigian S, Hopkins R, Goeree R. Surrogate outcomes: experiences at the Common Drug Review. *Cost Eff Resour Alloc*. 2013;11(1):31.
5. Kovic B, Jin X, Kennedy SA, et al. Evaluating progression-free survival as a surrogate outcome for health-related quality of life in oncology: A systematic review and quantitative analysis. *JAMA Intern Med*. 2018;178(12):1586-1596.
6. Booth CM, Eisenhauer EA. Progression-free survival: meaningful or simply measurable? *J Clin Oncol*. 2012;30(10):1030-1033.
7. Gutman SI, Piper M, Grant MD, et al. *Progression-free survival: What does it mean for psychological well-being or quality of life?* Rockville: Agency for Healthcare Research and Quality; 2013.
8. Haslam A, Hey S, Gill J, Prasad V. A systematic review of trial-level meta-analyses measuring the strength of association between surrogate end-points and overall survival in oncology. *Eur J Cancer*. 2019;106:196-211.
9. Canadian Agency for Drugs and Technologies in Health. *Procedure and submission guidelines for the CADTH Common Drug Review*. Ottawa: CADTH; 2019.
10. Angelis A, Lange A, Kanavos P. Using health technology assessment to assess the value of new medicines: results of a systematic review and expert consultation across eight European countries. *Eur J Health Econ*. 2018;19(1):123-152.
11. Pavlovic M. Challenges for relative effectiveness assessment and early access of cancer immunotherapies in Europe. *Front Med (Lausanne)*. 2016;3:56.
12. Jonsson B, Martinalbo J, Pignatti F. European Medicines Agency perspective on oncology study design for marketing authorization and beyond. *Clin Pharmacol Ther*. 2017;101(5):577-579.

13. Garrido MV, Mangiapane S. Surrogate outcomes in health technology assessment: an international comparison. *Int J Technol Assess Health Care*. 2009;25(03):315-22.
14. Elston J, Taylor RS. Use of surrogate outcomes in cost-effectiveness models: a review of United Kingdom health technology assessment reports. *Int J Technol Assess Health Care*. 2009;25(01):6-13.
15. Kleijnen S, Lipska I, Leonardo Alves T et al. Relative effectiveness assessments of oncology medicines for pricing and reimbursement decisions in European countries. *Ann Oncol*. 2016;27(9):1768-75.
16. Dakin H, Devlin N, Feng Y, Rice N, O'Neill P, Parkin D. The influence of cost-effectiveness and other factors on NICE decisions. *Health Econ*. 2015;24(10):1256-1271.
17. McCabe C, Claxton K, Culyer AJ. The NICE cost-effectiveness threshold: what it is and what that means. *Pharmacoeconomics*. 2008;26(9):733-744.
18. Rocchi A, Menon D, Verma S, Miller E. The role of economic evidence in Canadian oncology reimbursement decision-making: to lambda and beyond. *Value Health*. 2008;11(4):771-783.
19. Griffiths E, Vadlamudi N. CADTH's $50,000 cost-effectiveness threshold: fact or fiction? *Value Health*. 2016;19(7):A488-A489.
20. Davis C, Naci H, Gurpinar E, Poplavska E, Pinto A, Aggarwal A. Availability of evidence on overall survival and quality of life benefits of cancer drugs approved by the European Medicines Agency: A retrospective cohort study of drug approvals from 2009-2013. *BMJ*. 2017;359:j4530.
21. European Medicines Agency. *Guideline on the evaluation of anticancer medicinal products in man*. London: EMA; 2017.
22. Lavallée LT, Montori VM, Canfield SE, Breau RH. Advanced topics in evidence-based urologic oncology: surrogate endpoints. *Urol Oncol*. 2011;29(4):447-453.
23. Kreeftmeijer J, Ryan J, Van Engen A, Heemstra L. Hierarchy of clinical endpoints in HTA decision making in oncology. *Value Health*. 2015;18(3):A220.
24. Clement F, Harris A, Li J, Yong K, Lee K, Manns B. Using effectiveness and cost-effectiveness to make drug coverage decisions: a comparison of Britain, Australia and Canada. *JAMA*. 2009;302(13):1437-1443.
25. McCabe C, Bergmann L, Bosanquet N et al. Market and patient access to new oncology products in Europe: a current, multidisciplinary perspective. *Ann Oncol*. 2009;20(3):403-412.

26. Lim CS, Lee YG, Koh Y, Heo DS. International comparison of the factors influencing reimbursement of targeted anti-cancer drugs. *BMC Health Serv Res*. 2014;14:595.
27. The Council of the Federation. *The pan-Canadian Pharmaceutical Alliance*. Ottawa: Canada's Premiers; 2018.
28. National Institute for Health and Care Excellence. *Procedure for the review of Patient Access Scheme proposals*. London: Patient Access Scheme Liaison Unit; 2018.
29. National Institute for Health and Care Excellence. *Technology appraisal and highly specialised technologies appeals*. London: NICE; 2018.
30. Canadian Agency for Drugs and Technologies in Health. *CADTH pan-Canadian Oncology Drug Review*. Ottawa: CADTH; 2019.
31. Potter A, Knight A. The Cancer Drugs Fund. The Lancet UK Policy Matters. https://ukpolicymatters.thelancet.com/policy-summary-the-cancer-drugs-fund/. Published 2011. Accessed July 19, 2019.
32. National Health Service England. *Appraisal and funding of cancer drugs from July 2016 (including the new Cancer Drugs Fund)*. London: NHS England; 2016.
33. Cherny N, Dafni U, Bogaerts J et al. ESMO-Magnitude of Clinical Benefit Scale version 1.1. *Annals of Oncology*. 2017;28(10):2340-2366.
34. Schnipper L, Davidson N, Wollins D et al. Updating the American Society of Clinical Oncology value framework: Revisions and reflections in response to comments received. *J Clin Oncol*. 2016;34(24):2925-2934.
35. Sherrill B, Kaye JA, Sandin R, Cappelleri JC, Chen C. Review of meta-analyses evaluating surrogate endpoints for overall survival in oncology. *Onco Targets Ther*. 2012;5:287-296.
36. Tsoi B, Masucci L, Campbell K, Drummon M, O'Reilly D, Goeree R. Harmonization of reimbursement and regulatory approval processes: a systematic review of international experiences. *Expert Rev Pharmacoecon Outcomes Res*. 2013;13(4):497–511.