CURRENT
ONCOLOGY

Impact of country-specific EQ-5D-3L tariffs on the economic value of systemic therapies used in the treatment of metastatic pancreatic cancer

K. Lien BSc,* V.C. Tam BSc(Hon) MD,[†] Y.J. Ko MD MMSc SM,[†] N. Mittmann PhD,[‡] M.C. Cheung MD SM,[†] and K.K.W. Chan MD MSc[†§]

ABSTRACT

Background   Previous Canadian cost-effectiveness analyses in cancer based on the EQ-5D-3L (EuroQoL, Rotterdam, Netherlands) have commonly used U.K. or U.S. tariffs because the Canadian equivalent only just recently became available. The implications of using non-Canadian tariffs to inform decision-making are unclear. We aimed to re-evaluate an earlier cost-effectiveness analysis of therapies for metastatic pancreatic cancer (originally performed using U.S. tariffs) with tariffs from Canada and various other countries to determine the impact of using non-country-specific tariffs.

Methods   We used tariffs from Canada, the United States, the United Kingdom, Denmark, France, Germany, Japan, the Netherlands, and Spain to derive EQ-5D-3L utilities for the 10 health states in the pancreatic cancer model. Quality-adjusted life years (QALYs) and incremental cost-effectiveness ratios (ICERs) were generated, and probabilistic sensitivity analyses (PSAs) were performed.

Results   Canadian utilities are generally lower than the corresponding U.S. utilities and higher than those for the United Kingdom. Compared with the Canadian-valued scenarios, U.S. and U.K. estimates were statistically different for 3 and 9 scenarios respectively. Overall, 35% of the non-Canadian utilities (28 of 80) were significantly different, clinically, from the Canadian values. Canadian QALYs were 6% lower than those for the United States and 6% higher than those for the United Kingdom. When comparing the QALYs of each treatment with those of gemcitabine alone, the average percent change was +6.8% for a U.S. scenario and –7.5% for a U.K. scenario compared with a Canadian scenario. Consequently, Canadian ICERs were approximately 5.4% greater than those for the United States and 8.6% lower than those for the United Kingdom. Based on the PSAs and compared with the Canadian threshold value, the minimum willingness-to-pay threshold at which the combination chemotherapy regimen of gemcitabine–capecitabine is the most cost-effective is $5,239 less than in the United States and $11,986 more than in the United Kingdom.

Conclusions   The use of non-country-specific tariffs leads to significant differences in the derived utilities, ICERs, and PSA results. Past Canadian EQ-5D-3L–based cost-effectiveness analyses and related funding decisions might need to be re-visited using Canadian tariffs.

Key Words   Pancreatic cancer, cost-effectiveness, EQ-5D-3L, country-specific tariffs, weights

INTRODUCTION

In light of increasing health care costs, administrators and governments often consider cost-effectiveness analyses to be essential tools for making health care decisions and allocating resources[1]. The analyses compare the costs and outcomes of current medical interventions with alternative strategies, which can include combination therapy, the use of novel therapeutic agents, or discontinuation of treatment[2]. The desired outcome or effectiveness of a treatment is considered to be an increase in the quality of life or the life expectancy of a patient (or both), commonly

Correspondence to: Kelvin K.W. Chan, Sunnybrook Odette Cancer Centre, 2075 Bayview Avenue, T-Wing, Toronto, Ontario M4N 3M5.
E-mail: kelvin.chan@sunnybrook.ca ■ **DOI:** http://dx.doi.org/10.3747/co.22.2592

pooled as a single measurement of quality-adjusted life years (QALYs). Calculating the cost difference between the two interventions and dividing by the difference in QALYs yields the incremental cost-effectiveness ratio (ICER), the major outcome reported for such analyses[3]. Authorities such as the U.K. National Institute for Health and Clinical Excellence[4] and the Canadian Agency for Drugs and Technologies in Health[5] generally accept the use of QALYs. Cost–utility assessments in the field of oncology also frequently use QALYs[6].

Health utilities are often measured using instruments such as the Health Utilities Index (Health Utilities, Hamilton, ON) and the EQ-5D-3L (EuroQoL, Rotterdam, Netherlands)[7,8]. Those tools rank utility on a scale of 0–1 (1 representing the best possible health state and 0 representing death (that is, less than 0)[9].

The EQ-5D-3L in particular is regularly applied in health utility studies of cancer, and it is the endorsed instrument of the U.K. National Institute for Health and Clinical Excellence[10]. This multi-attribute utility instrument in the form of a patient-reported questionnaire represents a single health state by documenting 5 domains of health (mobility, self-care, usual activities, pain and discomfort, and anxiety and depression). Each domain is ranked on a scale of 1–3, representing none, some, or extreme problems in that area. Thus, 243 health states are possible.

To account for the loss of individual patient preferences, valuation studies have been used to create country-specific tariffs for the conversion of EQ-5D-3L scores into utilities[11]. Valuation studies are generally conducted as large population-based surveys that use time trade-off techniques to map a selected subset of EQ-5D-3L states to a scoring algorithm. The scoring algorithm establishes a statistical relationship between the utilities and the EQ-5D-3L scores to generate an index of country-specific tariffs[12], with the U.K. index being the most established. If a country-specific index is not available, the U.K. index is often used[13–15]. However, a number of studies have demonstrated that using a non-country-specific tariff is inappropriate because of variations in methodologies and cultures.

Concerns with methodologies include the fact that, in a few studies, survey populations differed, as did the variables included in the final modelling equation. For example, certain valuation studies surveyed the general public[16,17]; others surveyed patients[18] or caregivers[19] (or both). The variables included in modelling equations provide numeric values that account for the loss of utility that occurs when a dimension is not given a rating of 1. For example, the model might account for a mobility rating of 2 by subtracting 0.05 from the overall utility of the health state. A study that compared the modelling methodologies of several EQ-5D-3L valuation studies suggests that, although certain countries (Denmark, Germany, and the Netherlands, for example) generated similar, but higher, preference scores than those generated for the United Kingdom, others (such as the Spanish model) showed significantly more variance. These country-specific differences in utilities were also observed by other authors[20,21].

Differences in age, sex, standard of living, and preferences by health state in the general population of various countries can potentially contribute to cultural differences in a model[22,23]. For example, a French valuation study found that French respondents weighted problems in the mobility, self-care, and usual activity parameters more heavily than did U.K. respondents[24]. A comparison of standardized EQ-5D-3L data across countries found that prior living standards (per-capita gross domestic product) correlated the most with the EQ visual analogue scale scores (Spearman rank correlation: 0.58). Linear regression analysis showed that gross domestic product explained 67% of the EQ visual analogue scale ($p = 0.02$) when outliers were excluded and 29% when the outliers were included; per-capita health expenditure explained 26% of the mean visual analogue scale ($p = 0.03$). Finally, in a preliminary study of the relationships between national culture and EQ-5D value sets, power distance (that is, a society in which people expect and desire inequalities among themselves) and individualism were found to have moderate and strong correlations with pain and discomfort, and anxiety and depression[22].

Differences in methodology and culture can lead to differences in QALY calculations and cost-effectiveness analyses. Karlsson *et al.*[25] found that applying various national EQ-5D-3L tariffs to the same data can result in substantially different incremental QALY estimates. In their study, U.K. values were approximately 1.5 times those calculated using U.S. or Danish tariffs. One cost-effectiveness analysis conducted in the United States ($n = 301$) that compared treatments for Parkinson disease showed that the ICER calculated using U.S.-specific tariffs was higher than the ICER calculated using U.K. tariffs ($108,498/QALY vs. $42,989/QALY)[26].

Canada has a predominantly public health system, and so it is crucial that its limited resources are properly allocated. Hospital expenditures for cancer drugs are high; in 2009 alone, provincial cancer agencies were estimated to have spent $800 million[27]. Cost-effectiveness analyses can be used to influence funding decisions and to limit wasteful spending on cost-ineffective treatments.

Many Canadian cost-effectiveness analyses in cancer have been conducted using the utility scores derived from the U.K. tariffs[28–31] and, to a lesser extent, the U.S. tariffs[32] obtained from the EQ-5D-3L. The implications of this practice of using non-Canadian tariffs have not been determined, and the Canadian Agency for Drugs and Technologies in Health has not yet made an official recommendation on this topic. The previously mentioned studies conducted in the United States and Denmark suggest the possibility that differences in country-specific tariffs could lead to differences in the associated cost-effectiveness results, which might in turn alter reimbursement decisions. It is important to note that no currently published studies have delineated the effect that the use of Canada-specific tariffs (compared with tariffs from other countries, especially the United Kingdom and the United States) would have on cost-effectiveness analyses. In any given evaluation, adopting the appropriate tariffs is therefore crucial if the aim is to ensure that appropriate country-specific decisions are made.

A Canadian valuation of EQ-5D-3L tariffs was recently published[17], and its findings suggest that a review of cost-effectiveness analyses based in Canada might be necessary.

In the present study, we set out to determine whether the use of country-specific tariffs produces biased results in cost-effectiveness analyses. As an example, we used an earlier analysis that aimed to assess the cost-effectiveness of various therapies for metastatic pancreatic cancer from a Canadian perspective. That study was originally performed using U.S.-based preference tariffs; in the present analysis, it was updated to use alternative country-specific tariffs[32].

METHODS

We recently published a cost-effectiveness analysis for systemic therapies in pancreatic cancer from a Canadian public payer perspective[32]. In brief, a Markov analytic decision model was used to analyze a hypothetical cohort of patients with metastatic pancreatic cancer treated with one of four chemotherapy regimens. In the analysis, gemcitabine (GEM) was compared with three other combination therapies: gemcitabine–capecitabine (GEM-CAP), gemcitabine–erlotinib (GEM-ERL) and oxaliplatin, irinotecan, 5-fluorouracil, and folinic acid (FOLFIRINOX). Full treatment details, adverse events, and resource utilization can be found in the original publication.

The original study created 8 scenarios describing the symptoms of patients undergoing one of the four chemotherapy treatments. The symptoms included nausea and vomiting, diarrhea, hand–foot syndrome, stomatitis, febrile neutropenia, fatigue, rash, and neutropenia. One additional scenario was used to describe stable disease, and another, to designate a clinical scenario with only basic supportive care. For the original study, those scenarios were sent in the form of a survey to 60 medical oncologists in Canada who were requested to use the EQ-5D-3L to report their perception of the patient's health state. The oncologists chosen were all experts in noncolorectal gastrointestinal cancers and were familiar with the relevant treatments and side effects. Additional questions to ascertain respondent demographics were included in the survey, and the demographics of the oncologists who responded are presented in Table I here. More information on methods are provided in the original publication[32].

The EQ-5D-3L survey responses were converted into utility scores using each of the country-specific tariffs. These countries and regions were used: the United Kingdom[15], the United States[16], Canada[17], Denmark[33], France[24], Germany[34], Japan[35], the Netherlands[36], and Spain[37]. Each country's original valuation study developed its own mathematical model for determining country-specific tariffs. Those tariffs were used to score the EQ-5D-3L results obtained in our survey, generating a set of country-specific utilities for each individual responder and for each scenario. The utilities were then averaged to obtain the mean country-specific utilities for each scenario. Table II lists the model types according to each country's valuation study. For each scenario, the utility scores thus generated were compared with the Canadian equivalents using paired two-tailed t-tests; differences were considered significant at $p < 0.05$. To account for multiple comparisons, the Bonferroni correction was used. Furthermore, to determine the effects of utility differences on clinical practice, one-sample one-sided t-tests were performed on each of the

TABLE I Demographics of survey respondents

Characteristic	Respondents (*n* of 33)
Age	
30–39 Years	11
40–49 Years	16
50–59 Years	6
Sex	
Women	11
Men	22
Type of facility	
Academic centre	20
Academic centre and community cancer centre	1
Academic hospital	3
Community cancer centre	6
Community hospital	3
Years since residency or fellowship	
0–4	6
5–9	13
10–14	5
15–20	6
20–24	1
25–29	2
Familiarity with economic evaluations	
Not familiar	4
Somewhat familiar	21
Very familiar	8

10 scenarios using the differences in the country-specific utilities compared with the Canadian utilities. The null hypothesis was that the absolute difference in utilities was equal to or less than 0.06, the minimally important difference (MID) based on the literature[38].

Cost-effectiveness was determined as Canadian dollars per QALY for each of the country-specific utilities. The ICERs were generated using GEM alone as the base case. Probabilistic sensitivity analysis was performed for 10,000 simulations to determine the cost-effectiveness and net monetary benefit of each regimen over a range of willingness-to-pay (WTP) thresholds, expressed as

$$E \times WTP - C,$$

where E is effectiveness, WTP is willingness to pay, and C is cost.

RESULTS

Figure 1 shows the utility values obtained from the survey when Canadian, U.S., and U.K. tariffs were used, with Canadian values being the common comparator. Correlations of

TABLE II Background characteristics for time trade-off elicited country valuation studies

Country	Model type	Individuals analyzed (*n*)	Regression coefficients (*n*)	R^2	Reference
Canada	Random effects	1145	$10 + \alpha$	0.40	Bansback *et al.,* 2012[17]
United States	Random effects	3773	$10 + D1, I2^2, I3, I3^2$	0.38	Shaw *et al.,* 2005[16]
United Kingdom	Random effects	2997	$10 + \alpha + N3$	0.46	Dolan, 1997[15]
Denmark	Random effects	1332	$10 + \alpha$	0.66	Wittrup-Jensen *et al.,* 2009[33]
France	Random effects	443	$10 + N3$	NA	Chevalier *et al.,* 2013[24]
Germany	Random effects	334	$6 + \alpha + N3$	0.51	Greiner *et al.,* 2005[34]
Japan	Ordinary least squares	543	$10 + \alpha$	0.40	Tsuchiya *et al.,* 2002[35]
Netherlands	Random effects	298	$10 + \alpha + N3$	0.38	Lamers *et al.,* 2006[36]
Spain	Random effects	975	$10 + \alpha + N3$	0.60	Badia *et al.,* 2001[37]

α = intercept [indicator for any movement away from perfect health (11111)]; D1 = number of dimensions at level 2 or level 3 beyond the first; I2 = number of dimensions at level 2 beyond the first; I3 = number of dimensions at level 3 beyond the first; N3 = level 3 occurs within at least one dimension; NA = not available.
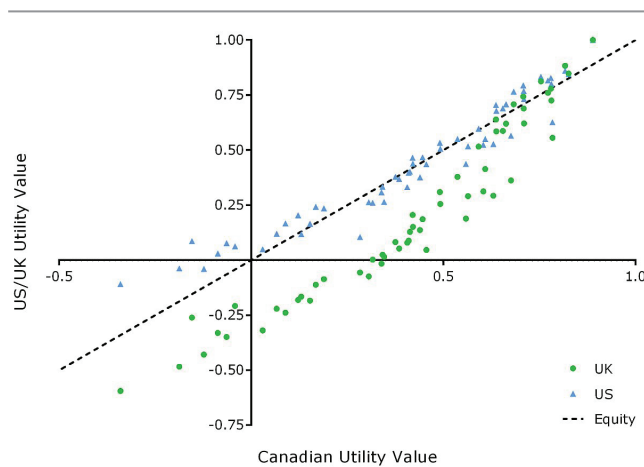

FIGURE 1 Canadian compared with U.S. and U.K. mean utilities.

the Canadian scores with those from the United States and United Kingdom were high: the R^2 using the Pearson product was 0.9361 and 0.9460 respectively. In general, U.S. utilities were greater than the corresponding Canadian ones; U.K. utilities were smaller. Table III summarizes the results of the EQ-5D-3L survey for all countries. Japan generally had the highest utility scores overall (7 of 10 scenarios); the Spanish and U.K. scores were generally the lowest (4 and 3 of 10 scenarios respectively). When comparing the utilities based on Canadian tariffs with the utilities based on tariffs from other countries for the 10 scenarios, statistical differences were observed for 10 utility values derived using Japan-based tariffs, 9 using U.K.- and Netherlands-based tariffs, 8 using Spain-based tariffs, 7 using France-based tariffs, 6 using Denmark-based tariffs, 3 using U.S.-based tariffs, and 2 using Germany-based tariffs. In terms of the MID, none of the differences between the Canada- and U.S.-derived utility scores were significant. On the other hand, 60% of the differences between Canada- and U.K.-derived utility scores were significantly greater than the MID, suggesting that differences in the U.K. model might lead

to clinically significant differences in utilities. After the U.S.-derived utilities, Germany-derived utilities are the most similar to Canadian utilities. Of the utility differences, 35% were statistically significantly greater than the MID.

Table IV shows the results of the cost-effectiveness analysis in terms of QALYs. No difference in the order of treatment efficacy was observed: FOLFIRINOX was consistently the most effective, followed by GEM-ERL, GEM-CAP, and GEM alone. Variation between countries in regimen QALYs was observed, ranging from a minimum difference of 0 to a maximum difference of 0.184. The U.S.-derived QALYs were 6% greater on average than the Canadian ones; U.K.-derived QALYs were 6% lower. Compared with GEM alone, the changes in QALYs for GEM-CAP, GEM-ERL, and FOLFIRINOX were 0.048, 0.076, and 0.214 when Canada-derived, and 0.052, 0.081, and 0.226 when U.S.-derived. Furthermore, when compared with Canadian-derived QALYs, the average change over all three treatments was 6.8% higher for U.S.-derived QALYs and 7.5% lower for U.K.-derived QALYs. Interestingly, although only 1 German scenario resulted in a significant utility difference larger than the MID (Table III), the expected QALYs for Germany were the most different when compared with the Canadian ones (>20% different on average). The QALYs based on Danish tariffs were the most similar to the Canadian QALYs.

Table V shows the results of the cost-effectiveness analysis in terms of ICERs, with results, as expected, following the same trend shown by the QALY values. Using Canadian tariffs and comparing combination treatment with GEM alone, the ICERs were $84,475 for GEM-CAP, $155,459 for GEM-ERL, and $130,670 for FOLFIRINOX. The equivalent ICERs generated using U.S. tariffs were approximately 5.4% lower; the use of U.K. tariffs led to an 8.6% average increase in the ICERs. The largest differences between country-specific ICERs compared with Canadian ones was 17% (German tariffs). The ICERs generated using tariffs from Denmark were the most similar to the Canadian ICERs (2% difference on average). Cost-effectiveness acceptability curves for U.K., U.S., and Canadian values (Figure 2) visually show the differences in treatment acceptability.

TABLE III   Mean utility scores obtained from the EQ-5D-3L survey

Scenario comparator	Score for ...									
	Nausea and vomiting	Diarrhea	Hand-foot syndrome	Stomatitis	Febrile neutropenia	Fatigue	Rash	Neuropathy	Stable disease	Supportive care
Country-specific tariff[a]										
Canada	0.498	0.497	0.393	0.236	0.579	0.214	0.603	0.485	0.687	0.048
United States[b]	0.526	0.508	0.409	0.279	0.586	0.243	**0.626**	0.494	**0.720**	**0.136**
United Kingdom	**0.352**	**0.328**	**0.179**	**−0.038**	**0.454**	**−0.053**	**0.487**	**0.320**	0.643	**−0.250**
Denmark	0.442	**0.444**	**0.317**	**0.097**	0.547	**0.071**	**0.574**	0.442	0.673	**−0.148**
France	**0.375**	**0.339**	**0.217**	**0.129**	**0.449**	0.145	**0.474**	**0.316**	0.676	−0.010
Germany	0.569	0.542	0.416	0.230	0.651	0.208	**0.689**	0.528	**0.812**	0.064
Japan	**0.572**	**0.568**	**0.493**	**0.384**	**0.649**	**0.287**	**0.688**	**0.575**	**0.788**	**0.221**
Netherlands	**0.415**	**0.398**	**0.271**	**0.083**	**0.491**	**0.112**	**0.507**	**0.388**	0.630	**−0.088**
Spain	**0.348**	**0.321**	**0.169**	**−0.024**	**0.470**	**−0.135**	0.542	**0.328**	0.695	**−0.243**
Canadian value[c]										
Canada	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
United States*	0.028	0.011	0.016	0.043	0.006	0.029	0.024	0.009	0.033	0.088
United Kingdom	−0.146	**−0.169**	**−0.213**	**−0.274**	−0.126	**−0.267**	−0.116	**−0.165**	−0.044	**−0.298**
Denmark	−0.056	−0.052	−0.075	**−0.139**	−0.032	**−0.143**	−0.029	−0.043	−0.014	**−0.197**
France	−0.123	**−0.158**	**−0.176**	−0.108	**−0.130**	−0.069	**−0.129**	**−0.169**	−0.011	−0.059
Germany	0.071	0.046	0.023	−0.006	0.072	−0.006	0.086	0.043	**0.125**	0.016
Japan	0.074	0.072	**0.100**	**0.147**	0.070	0.073	0.085	0.090	**0.101**	**0.173**
Netherlands	−0.083	−0.099	**−0.121**	**−0.153**	−0.088	−0.103	−0.095	−0.097	−0.057	**−0.137**
Spain	−0.150	**−0.176**	**−0.224**	**−0.261**	−0.109	**−0.349**	−0.061	**−0.157**	0.008	**−0.291**

[a]   Boldface type indicates values that, using the Bonferroni correction ($p < 0.000625$), are significantly different from Canadian values.
[b]   Values used in the original study.
[c]   Boldface type indicates values that are significantly larger than a minimally important difference of 0.06 ($p < 0.000625$)[38].

Table VI shows the probability of a treatment regimen being cost-effective at specified WTP thresholds. At a low WTP ($50,000), GEM had the highest probability of being cost-effective for all countries. At a WTP of $75,000, GEM remained the most cost-effective treatment for most of the countries (ranging from 60.57% for the United States to 82.12% for the United Kingdom); however, at that threshold, GEM-CAP is more effective when German or Japanese tariffs are used. When the WTP is $100,000, all country-specific tariffs result in treatment with GEM-CAP being the most likely to be cost-effective. However, probabilities range from 57.11% (United Kingdom) to 92.76% (Japan). Finally, at higher WTP thresholds (>$150,000), FOLFIRINOX is the treatment most likely to be cost-effective using any of the country indices. Figure 3 graphically summarizes the probabilities for U.K., Canadian, and U.S. valuations.

The WTP range at which GEM-CAP is the most cost-effective treatment also varied depending on country-specific tariffs (Table VI). Below the lower limits of the ranges, GEM is the most cost-effective, but FOLFIRINOX is the most cost-effective above the upper limits. At no WTP threshold is GEM-ERL cost-effective—a result that is consistent for all country valuations. The largest difference in the thresholds between which GEM-CAP becomes the

treatment most likely to be cost-effective is found when comparing German and U.K. indices, with the German threshold being $25,361 lower than that of the United Kingdom. For FOLFIRINOX, the largest difference is between Germany and France, with Germany's threshold being $32,185 less than that of France.

DISCUSSION AND CONCLUSIONS

The present study used Canadian-based tariffs derived from responses on the EQ-5D-3L tool to compare the results of a cost-effectiveness analysis of first-line systemic therapies in metastatic pancreatic cancer with results generated using tariffs from the United States, the United Kingdom, and various other countries. Comparisons of the country-specific indices showed significant differences between the valuations of the derived utility scores. Those differences potentially have interesting clinical applications. According to a study by Pickard *et al.*[38], the minimally important difference in EQ-5D-3L scores in cancer is estimated to be 0.06. Thus, we found that 35% of the differences in EQ-5D-3L scores from Canada and from other countries could be considered clinically different, providing evidence that using country-specific EQ-5D-3L

tariffs is important in ensuring that clinically meaningful differences are captured.

The differences we uncovered had an observable effect on the results of the Markov modelling simulation.

Compared with our original analysis (which used U.S. tariffs), the analysis using Canadian tariffs resulted in QALYs that were 6.8% lower for interventions compared with standard GEM. Tariffs from the United Kingdom,

TABLE IV Quality-adjusted life years (QALYs) for metastatic pancreatic cancer treatments based on country-specific EQ-5D-3L tariffs

Tariff country	QALY				ΔQALY compared with Gem			%ΔQALY compared with Canada		
	Gem	Gem–Cap	Gem–Erl	FOLFIRINOX	Gem–Cap	Gem–Erl	FOLFIRINOX	Gem–Cap	Gem–Erl	FOLFIRINOX
Canada	0.479	0.527	0.555	0.693	0.048	0.076	0.214	—	—	—
United States	0.508	0.560	0.589	0.734	0.052	0.081	0.226	5.94	5.55	5.73
United Kingdom	0.449	0.491	0.522	0.650	0.042	0.073	0.201	–13.64	–3.65	–5.76
Denmark	0.468	0.514	0.544	0.679	0.046	0.076	0.211	–4.62	0.027	–1.69
France	0.454	0.497	0.526	0.655	0.043	0.072	0.201	–11.53	–5.14	–6.07
Germany	0.577	0.634	0.669	0.834	0.057	0.092	0.257	18.01	20.32	18.56
Japan	0.535	0.589	0.619	0.773	0.054	0.084	0.238	11.62	9.87	11.21
Netherlands	0.452	0.496	0.524	0.654	0.044	0.072	0.202	–10.33	–5.26	–5.61
Spain	0.479	0.522	0.558	0.694	0.043	0.079	0.215	–10.42	3.95	0.59

Gem = single-agent gemcitabine; Gem-Cap = gemcitabine–capecitabine; Gem-Erl = gemcitabine–erlotinib; FOLFIRINOX = leucovorin–5-fluorouracil–irinotecan–oxaliplatin.

TABLE V Results of the cost-effectiveness analysis expressed using the incremental cost-effectiveness ratio (ICER)

Tariff country	ICER compared with Gem alone (CA$/QALYs)			%ΔICER compared with Canada		
	Gem-Cap	Gem-Erl	FOLFIRINOX	Gem-Cap	Gem-Erl	FOLFIRINOX
Canada	84,475	155,459	130,670	—	—	—
United States	79,735	147,285	123,590	–5.61	–5.26	–5.42
United Kingdom	97,816	161,344	138,651	15.79	3.79	6.11
Denmark	88,504	155,504	132,905	4.77	0.03	1.71
France	94,679	164,143	138,856	12.08	5.59	6.26
Germany	71,424	128,460	108,599	–15.45	–17.37	–16.89
Japan	75,392	140,694	117,269	–10.75	–9.50	–10.26
Netherlands	92,527	164,143	138,168	9.53	5.59	5.74
Spain	92,527	149,598	129,814	9.53	–3.77	–0.66

Gem = gemcitabine; Cap = capecitabine; Erl = erlotinib; FOLFIRINOX = leucovorin–5-fluorouracil–irinotecan–oxaliplatin.
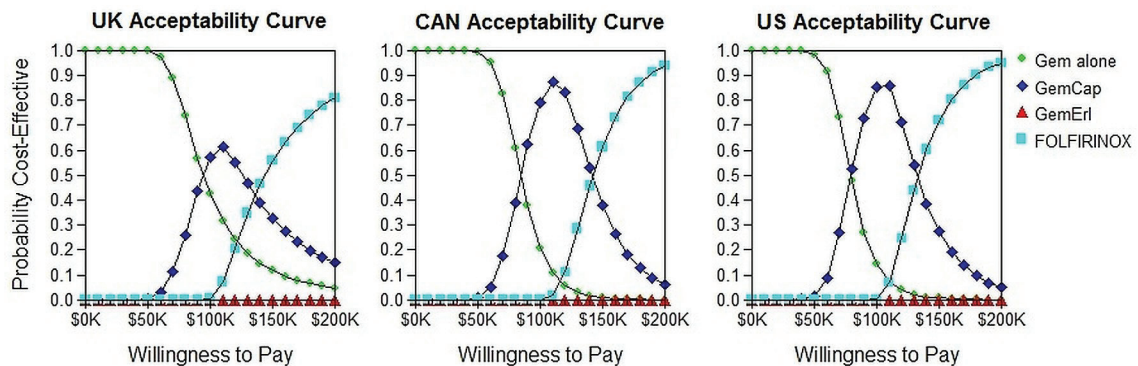

FIGURE 2 Cost-effectiveness acceptability curves at various willingness-to-pay thresholds for the United Kingdom, Canada (CAN), and the United States. Gem =gemcitabine; Cap = capecitabine; Erl = erlotinib; FOLFIRINOX = leucovorin–5-fluorouracil–irinotecan–oxaliplatin.

**TABLE VI** Percentage probability of achieving the highest net monetary benefit at specified willingness-to-pay (WTP) thresholds

Country	Treatment	Threshold (CA$)/QALY					
		0	50,000	75,000	100,000	150,000	200,000
Canada							
	Gem	100	99.19	72.55	20.92	0.81	0.04
	Gem–Cap	0	0.81	27.45	79.06	37.91	6.2
	Gem–Erl	0	0	0	0	0	0
	FOLFIRINOX	0	0	0	0.02	61.28	93.76
United States							
	Gem	100	98.46	60.57	14.24	0.74	0.02
	Gem–Cap	0	1.54	39.43	85.52	27.41	5.11
	Gem–Erl	0	0	0	0	0	0
	FOLFIRINOX	0	0	0	0.24	71.85	94.87
United Kingdom							
	Gem	100	99.6	82.12	42.45	11.48	3.98
	Gem–Cap	0	0.4	17.88	57.11	33.03	14.98
	Gem–Erl	0	0	0	0	0	0
	FOLFIRINOX	0	0	0	0.44	55.49	81.04
Denmark							
	Gem	100	99.42	76.86	31.12	3.22	0.37
	Gem–Cap	0	0.58	23.14	68.77	37.74	9.64
	Gem–Erl	0	0	0	0	0	0
	FOLFIRINOX	0	0	0	0.11	59.04	89.99
France							
	Gem	100	99.46	79	38.38	9.25	2.84
	Gem–Cap	0	0.54	21	61.38	36.82	16.94
	Gem–Erl	0	0	0	0	0	0
	FOLFIRINOX	0	0	0	0.24	53.93	80.22
Germany							
	Gem	100	95.98	36.1	8.68	1.55	0.34
	Gem–Cap	0	4.02	63.9	84.73	12.92	4.93
	Gem–Erl	0	0	0	0	0	0
	FOLFIRINOX	0	0	0	6.59	85.53	94.73
Japan							
	Gem	100	97.76	49.24	6.89	0.07	0.02
	Gem–Cap	0	2.24	50.76	92.76	17.52	2.06
	Gem–Erl	0	0	0	0	0	0
	FOLFIRINOX	0	0	0	0.35	82.41	97.92
Netherlands							
	Gem	100	99.29	78.89	37.93	7.96	2.1
	Gem–Cap	0	0.71	21.11	61.8	39.05	15.98
	Gem–Erl	0	0	0	0	0	0
	FOLFIRINOX	0	0	0	0.27	52.99	81.92
Spain							
	Gem	100	99.22	75.86	33.32	9.01	2.97
	Gem–Cap	0	0.78	24.14	65.64	24.92	11.02
	Gem–Erl	0	0	0	0	0	0
	FOLFIRINOX	0	0	0	1.04	66.07	86.01

TABLE VI   Continued

Country	Treatment	Threshold (CA$)/QALY					
		0	50,000	75,000	100,000	150,000	200,000
		Range of minimum WTP (CA$/QALY) at which Gem–Cap is most cost-effective[a]					
Canada				84,774–144,325			
United States				79,535–136,248			
United Kingdom				96,760–149,424			
Denmark				89,294–145,271			
France				93,559–151,520			
Germany				71,399–119,335			
Japan				74,888–130,132			
Netherlands				92,886–151,108			
Spain				90,952–139,358			

[a]   Based on probabilistic sensitivity analysis.
Gem = gemcitabine; Cap = capecitabine; Erl = erlotinib; FOLFIRINOX = leucovorin–5-fluorouracil–irinotecan–oxaliplatin.
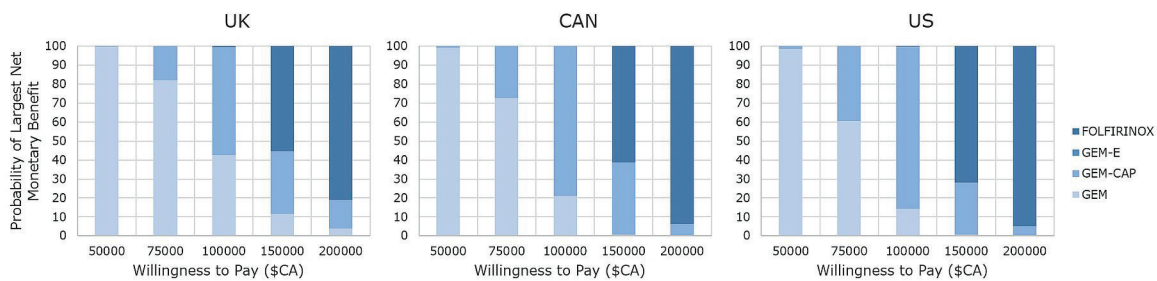
FIGURE 3   Probability of largest net monetary benefit at specified willingness-to-pay thresholds for the United Kingdom, Canada (CAN), and the United States [all values expressed in Canadian dollars ($CA)]. FOLFIRINOX = leucovorin–5-fluorouracil–irinotecan–oxaliplatin; GEM-E = gemcitabine–erlotinib; GEM-CAP = gemcitabine–capecitabine; GEM = single-agent gemcitabine.

which were often used as a surrogate index before Canadian tariffs were available, resulted in an expected change in QALYs that was, on average, 7.5% less than those calculated using Canadian tariffs. Those QALY differences contributed to variations in the associated ICERs and probabilities of the highest net monetary benefit at specific WTP thresholds in the final analysis. According to a survey of U.S. and Canadian oncologists, a common WTP range is $50,000–$100,000[39]. Between those thresholds, GEM is most likely to provide the highest net monetary benefit; however, the probability of that outcome is 64% in Canada, 58% in the United States, and 75% in the United Kingdom. At those intermediate WTP thresholds, the probability of treatments being the most cost-effective varies the most between countries and could have an effect on funding decisions.

Our study illustrates the effect of using non-country-specific tariffs on the utilities and results in a cost-effectiveness analysis in a pancreatic cancer model. Differences in country-specific EQ-5D-3L tariffs might be a result of methodologic variations between modelling studies. In obtaining EQ-5D-3L results, some authors have surveyed different populations (for example, a general population vs. patients and caregivers), leading to potential alternative perceptions of health. Most valuation studies attempt to fit multiple

models to their results so as to increase the R^2, leading to disparities in the number of variables included in the final models. For example, the N3 model used by most countries and the D1 model used in the United States could result in variations that are partly responsible for the observed differences in utilities. Sampling variability in each valuation study can also contribute to those differences.

Regardless of the source of the differences in tariffs between countries, the conventional way of applying the tariffs is to use the point estimate of each health state to calculate utilities, which ignores the sampling variation (or uncertainty) associated with the point estimate. However, because only one valuation study is usually performed in each country, there is no way to adequately determine if the failure to account for variability is the cause of differences in the utilities. To counteract that difficulty, Xie *et al.*[12] created a checklist for reporting valuation studies ("CREATE") that minimizes the heterogeneity caused by factors other than population health preferences. Items include the rationale for choosing the target population and the representativeness of the respondents with respect to the target population, among others. Future valuation studies should be standardized to implement CREATE, and more research into the causes of the differences in modelling valuations should be conducted.

One limitation of the present study is that, instead of patients, medical oncologists were surveyed to generate the EQ-5D-3L states for each hypothetical scenario. Some of the literature[40] suggests that it might be reasonable to use health care workers as surrogates for patients in obtaining health utilities, but this approach is controversial, and further research on the topic is needed. However, our original analysis used both one-way deterministic sensitivity analyses and probabilistic sensitivity analyses to examine the uncertainty with respect to utilities, finding that the results were robust to utility variability[32]. Furthermore, because the same values were used to calculate utilities according to each country's tariff (for example, 11232 was converted using each country's model), it is likely that any differences observed in the analysis were a result of significant differences in the valuation models rather than in the absolute 5-dimensional score. Another limitation of our study is that our survey generally described poor health states, which could amplify any differences between the models generated by the valuation studies. To validate our results, future studies should be expanded to analyze both poor and good health states.

We hope that the findings of our study will encourage other investigators and policymakers to revisit previous models and decisions that used non-country-specific tariffs. Canadian analyses that used U.K. tariffs were sensitive to utilities, and any studies that deemed that an intervention to be cost-ineffective should be reassessed because of the lower utility scores that would result from the use of U.K. instead of Canadian tariffs. Countries lacking their own valuation studies to determine EQ-5D-3L tariffs should strongly consider developing their own valuation studies rather than using tariffs from other countries when performing their cost-effectiveness analyses to inform policymaking. Given that most of the obvious differences in utilities appeared to occur for poor health states with low utilities (or even negative utilities), models describing health conditions involving poor health states—such as metastatic cancer or chronic severe cardiovascular diseases, including severe heart failure and stroke—are potentially more susceptible to this issue. Additionally, because of the rapidly rising costs of oncology drugs and increased reliance by many jurisdictions internationally on economic evaluations in drug funding decisions, our study assists oncologists—who are becoming more interested in learning the details of economic evaluations—by providing more knowledge so that they can engage in the discussion of drug assessment and funding.

CONFLICT OF INTEREST DISCLOSURES

AUTHOR AFFILIATIONS

*Odette Cancer Centre, Sunnybrook Health Sciences Centre, Toronto, ON; †Division of Medical Oncology and Hematology, Sunnybrook Odette Cancer Centre, University of Toronto, Toronto, ON; ‡Health Outcomes and Pharmacoeconomics Research Centre, Sunnybrook Health Sciences Centre, Toronto, ON; §Division of Biostatistics, Dalla Lana School of Public Health, University of Toronto, Toronto, ON.

REFERENCES

1. Russell LB, Gold MR, Siegel JE, Daniels N, Weinstein MC. The role of cost-effectiveness analysis in health and medicine. Panel on Cost-Effectiveness in Health and Medicine. *JAMA* 1996;276:1172–7.
2. Detsky AS, Naglie IG. A clinician's guide to cost-effectiveness analysis. *Ann Intern Med* 1990;113:147–54.
3. Rasanen P, Roine E, Sintonen H, Semberg-Konttinen V, Ryynanen OP, Roine R. Use of quality-adjusted life years for the estimation of effectiveness of health care: a systematic literature review. *Int J Technol Assess Health Care* 2006;22:235–41.
4. National Institute for Health and Care Excellence (NICE). *Social Value Judgements: Principles for the Development of NICE Guidance.* 2nd ed. London, UK: NICE; 2008.
5. Canadian Agency for Drugs and Technologies in Health (CADTH). *Guidelines for the Economic Evaluation of Health Technologies: Canada.* 3rd ed. Ottawa, ON: CADTH; 2006.
6. Earle CC, Chapman RH, Baker CS, *et al.* Systematic overview of cost–utility assessments in oncology. *J Clin Oncol* 2000;18:3302–17.
7. Furlong WJ, Feeny DH, Torrance GW, Barr RD. The Health Utilities Index (HUI) system for assessing health-related quality of life in clinical studies. *Ann Med* 2001;33:375–84.
8. EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy* 1990;16:199–208.
9. Brazier J, Ratcliffe J, Salomon JA, Tsuchiya A. *Measuring and Valuing Health Benefits for Economic Evaluation.* Oxford, UK: Oxford University Press; 2007.
10. Pickard AS, Wilke CT, Lin HW, Lloyd A. Health utilities using the EQ-5D in studies of cancer. *Pharmacoeconomics* 2007;25:365–84.
11. Szende A, Williams A, eds, on behalf of the EuroQol Group's International Task Force on Self-Reported Health. *Measuring Self-Reported Population Health: An International Perspective Based on EQ-5D.* Rotterdam, Netherlands: EuroQol; 2004.
12. Xie F, Gaebel K, Perampaladas K, Doble B, Pullenayegum E. Comparing EQ-5D valuation studies: a systematic review and methodological reporting checklist. *Med Decis Making* 2014;34:8–20.
13. Noyes K, Dick AW, Holloway RG on behalf of the Parkinson Study Group. Pramipexole versus levodopa in patients with early Parkinson's disease: effect on generic and disease-specific quality of life. *Value Health* 2006;9:28–38.
14. Hung MC, Lu HM, Chen L, *et al.* Cost per QALY (quality-adjusted life year) and lifetime cost of prolonged mechanical ventilation in Taiwan. *PLoS One* 2012;7:e44043.
15. Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;35:1095–108.
16. Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. *Med Care* 2005;43:203–20.
17. Bansback N, Tsuchiya A, Brazier J, Anis A. Canadian valuation of EQ-5D health states: preliminary value set and considerations for future valuation studies. *PLoS One* 2012;7:e31115.
18. Augustovski FA, Irazola VE, Velazquez AP, Gibbons L, Craig BM. Argentine valuation of the EQ-5D health states. *Value Health* 2009;12:587–96.
19. Yusof FA, Goh A, Azmi S. Estimating an EQ-5D value set for Malaysia using time trade-off and visual analogue scale methods. *Value Health* 2012;15(suppl):S85–90.
20. Norman R, Cronin P, Viney R, King M, Street D, Ratcliffe J. International comparisons in valuing EQ-5D health states: a review and analysis. *Value Health* 2009;12:1194–200.

21. Knies S, Evers SM, Candel MJ, Severens JL, Ament AJ. Utilities of the EQ-5D: transferable or not? *Pharmacoeconomics* 2009;27:767–79.
22. Bailey H, Kind P. Preliminary findings of an investigation into the relationship between national culture and EQ-5D value sets. *Qual Life Res* 2010;19:1145–54.
23. Szende A, Janssen B, Cabasés J, eds. *Self-Reported Population Health: An International Perspective Based on EQ-5D.* Dordrecht, Netherlands: Springer Netherlands; 2014.
24. Chevalier J, de Pouvourville G. Valuing EQ-5D using time trade-off in France. *Eur J Health Econ* 2013;14:57–66.
25. Karlsson JA, Nilsson JA, Neovius M, *et al.* National EQ-5D tariffs and quality-adjusted life-year estimation: comparison of UK, US and Danish utilities in south Swedish rheumatoid arthritis patients. *Ann Rheum Dis* 2011;70:2163–6.
26. Noyes K, Dick AW, Holloway RG. The implications of using US-specific EQ-5D preference weights for cost-effectiveness evaluation. *Med Decis Making* 2007;27:327–34.
27. Canadian Institute for Health Information (cihi). *Drivers of Prescription Drug Spending in Canada.* Ottawa, ON: cihi; 2012.
28. Ewara EM, Zaric GS, Welch S, Sarma S. Cost-effectiveness of first-line treatments for patients with *KRAS* wild-type metastatic colorectal cancer. *Curr Oncol* 2014;21:e541–50.
29. Delea TE, Amdahl J, Nakhaipour HR, *et al.* Cost-effectiveness of pazopanib in advanced soft-tissue sarcoma in Canada. *Curr Oncol* 2014;21:e748–59.
30. Delea TE, El Ouagari K, Rotter J, Wang A, Kaura S, Morgan GJ. Cost-effectiveness of zoledronic acid compared with clodronate in multiple myeloma. *Curr Oncol* 2012;19:e392–403.
31. Attard CL, Brown S, Alloul K, Moore MJ. Cost-effectiveness of folfirinox for first-line treatment of metastatic pancreatic cancer. *Curr Oncol* 2014;21:e41–51.
32. Tam VC, Ko YJ, Mittmann N, *et al.* Cost-effectiveness of systemic therapies for metastatic pancreatic cancer. *Curr Oncol* 2013;20:e90–e106.
33. Wittrup-Jensen KU, Lauridsen J, Gudex C, Pedersen KM. Generation of a Danish tto value set for EQ-5D health states. *Scand J Public Health* 2009;37:459–66.
34. Greiner W, Claes C, Busschbach JJ, von der Schulenburg JM. Validating the EQ-5D with time trade off for the German population. *Eur J Health Eco* 2005;6:124–30.
35. Tsuchiya A, Ikeda S, Ikegami N, *et al.* Estimating an EQ-5D population value set: the case of Japan. *Health Econ* 2002;11:341–53.
36. Lamers LM, McDonnell J, Stalmeier PF, Krabbe PF, Busschbach JJ. The Dutch tariff: results and arguments for an effective design for national EQ-5D valuation studies. *Health Econ* 2006;15:1121–32.
37. Badia X, Roset M, Herdman M, Kind P. A comparison of United Kingdom and Spanish general population time trade-off values for EQ-5D health states. *Med Decis Making* 2001;21:7–16.
38. Pickard AS, Neary MP, Cella D. Estimation of minimally important differences in EQ-5D utility and vas scores in cancer. *Health Qual Life Outcomes* 2007;5:70.
39. Berry SR, Bell CM, Ubel PA, *et al.* Continental Divide? The attitudes of US and Canadian oncologists on the costs, cost-effectiveness, and health policies associated with new cancer drugs. *J Clin Oncol* 2010;28:4149–53.
40. Suarez-Almazor ME, Conner-Spady B. Rating of arthritis health states by patients, physicians, and the general public. Implications for cost-utility analyses. *J Rheumatol* 2001;28:648–56.