ScienceDirect

Contents lists available at sciencedirect.com
Journal homepage: www.elsevier.com/locate/jval

Health Policy Analysis

Health Technology Assessment–Informed Decision Making by the Federal Joint Committee/Institute for Quality and Efficiency in Health Care in Germany and the National Institute for Health and Care Excellence in England: The Role of Budget Impact

Ramon Schaefer, MA, BSc, Diego Hernández, PhD, Till Bärnighausen, MD, PhD, ScD, Peter Kolominsky-Rabas, MD, PhD, MBA, Michael Schlander, MD, PhD, MBA

ABSTRACT

Objectives: This study aimed to test (official) evaluation criteria including the potential role of budget impact (BI) on health technology assessment (HTA) outcomes published by the Federal Joint Committee (Gemeinsamer Bundesausschuss [GBA]) and the Institute for Quality and Efficiency in Health Care (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen [IQWiG]) in Germany as well as the National Institute for Health and Care Excellence (NICE) in England.

Methods: Data were extracted from all publicly available GBA decisions and IQWiG assessments as well as NICE single technology appraisals between January 2011 and June 2018, and information with regard to evaluation criteria used by these agencies was collected. Data were analyzed using logistic regression to estimate the effect of the BI on the HTA outcomes while controlling for criteria used by GBA/IQWiG and NICE.

Results: NICE recommendations are largely driven by the incremental cost-effectiveness ratio and, if applicable, by end-of-life criteria ($P < .01$). While IQWiG assessments are significantly affected by the availability of randomized controlled trials and patient-relevant endpoints ($P < .01$), GBA appraisals primarily focus on endpoints ($P < .01$). The BI correlated with NICE single technology appraisals (inverted-U relationship, $P < .1$) and IQWiG recommendations (increasing linear relationship, $P < .05$), but not with GBA decisions ($P > .1$). Nevertheless, given that IQWiG assessments seem to be more rigorous than GBA appraisals regarding the consideration of evidence-based evaluation criteria, decisions by GBA might be negatively associated with the BI.

Conclusions: Results reveal that GBA/IQWiG and NICE follow their official evaluation criteria consistently. After controlling for all significant variables, the BI seems to have an (independent) effect on HTA outcomes as well.

Keywords: budget impact, evaluation criteria, Federal Joint Committee, health technology assessment, Institute for Quality and Efficiency in Health Care, National Institute for Health and Care Excellence.

VALUE HEALTH. 2023; 26(7):1032–1044

Introduction

The introduction of health technology assessments (HTAs) for the systematic evaluation of medical interventions can be traced back to the emerging need for an effective and efficient use of innovative health technologies.[1,2] HTAs have been described as a multidisciplinary process to support decision making,[1-3] but their practical implementation rests primarily on 2 pillars: the principles of evidence-based medicine (EBM) for the assessment of clinical effectiveness and an evaluation of efficiency—usually by means of a variant of cost-effectiveness (CE) analysis (CEA).

Under the premise that CEA focuses on cost per patient or intervention, international guidelines[4-6] suggest that budget impact (BI) analyses (BIAs) might play a greater role in HTA. In fact, BIAs are discussed as a complementary component of a comprehensive economic evaluation to estimate the potential impact of a new health technology on the available budget (or overall costs). The focus of BIAs is on assessing the "affordability" of an intervention for a health system, and thus, they should be made for an adequate time horizon from the perspective of payers and decision makers.[5,6]

Internationally recognized HTA agencies such as the National Institute for Health and Care Excellence (NICE) in England and the Federal Joint Committee (Gemeinsamer Bundesausschuss [GBA]) and the Institute for Quality and Efficiency in Health Care (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen [IQWiG]) in Germany require pharmaceutical companies to conduct BIAs when submitting documents for the evaluation of new health technologies. Interestingly, both GBA/IQWiG[7] and NICE[8] do not formally consider BI estimations in HTA recommendations, and therefore, its actual role remains unclear.

NICE published information on the financial impact of technologies recommended for routine use within the UK's National Health Service (NHS) in so-called Costing Statements.[9] Since April 2017, a BI test has been introduced to assess the financial impact of a new technology over the first 3 years of its use in the NHS.[8] For technologies exceeding a BI of £20 million in any of the first 3 years, NICE will trigger commercial negotiations between the NHS and the pharmaceutical company, which may limit patient access to some innovative treatments.[10]

Although a number of previous studies[11-15] have analyzed assessment criteria used in NICE guidance, only 2 empirical articles reported results on the BI.[12,16] Dakin et al[12] found no significant impact of the BI in their analysis; nevertheless, interventions recommended for restricted (or optimized) use had a significantly higher BI than those recommended. A more recent study by Mauskopf et al[16] indicated a significant correlation between the BI and the level of reimbursement restrictions recommended by NICE, after controlling for clinical and CE.

In England and Wales, NICE guidance on the use of existing and new technologies including rapid reviews for single indications (single technology appraisals [STAs]) are legally binding within the NHS.[8] NICE recommendations are usually based on a review of clinical and economic evidence as well as additional factors related to ethical or social values.[17] Nevertheless, the application of a cost per quality-adjusted life-year (QALY) threshold (£20 000-30 000/QALY) suggests that NICE recommendations heavily rely on incremental CE ratios (ICERs) as an indicator of value for money.[11-15] Given controversies against the background of an increase of negative recommendations in (primarily) cancer-related STAs, criteria for end-of-life (EoL) treatments were implemented.[10,18] Later, the Cancer Drugs Fund (CDF) was introduced to provide timely restricted funding for technologies failing to meet EoL criteria.[19]

In Germany, since the enactment of the Pharmaceutical Market Restructuring Act (Arzneimittelmarktneuordnungsgesetz [AMNOG]), early benefit assessments (EBAs) of newly authorized drugs were officially introduced, and pharmaceutical manufacturers are required to submit a detailed value dossier including estimations on the overall cost impact of a new therapy.[20] To date, no empirical data have been published on the role of the BI within the German HTA context. Although one article by Fischer and Stargardt[21] identified a significant association of positive GBA decisions and higher annual treatment cost per patient, no clear pattern seemed to be existent for the BI. The authors found that both annual treatment cost per patient and the maximum possible BI were significantly higher in manufacturer dossiers than GBA decisions.[21]

In a 2-stage assessment procedure, GBA usually commissions IQWiG with the assessment of the (clinical) evidence submitted by a pharmaceutical company. Based on IQWiG recommendations, GBA makes the final decision in terms of the extent and certainty of the additional clinical benefit. This is different for drugs with an orphan designation, because they are legally assumed to confer some added benefit as long as actual sales do not exceed social health insurance (SHI) expenditures of €50 million per year.[7] According to GBA procedures[7] and IQWiG methods,[22] EBAs primarily focus on comparative effectiveness based on the robust principles of EBM. In contrast to the evaluation process adopted by NICE, both GBA and IQWiG rejected CEAs and the incremental cost per QALY metric. Instead, IQWiG derived the efficiency frontier approach for optional economic evaluation of health interventions, but has not (yet) applied it in the context of EBAs.[22]

Against this background, our study aimed to extend upon the existing HTA literature regarding (official) evaluation criteria, explore variations in HTA outcomes, and provide insights on the potential role of the BI in decision making for both GBA/IQWiG and NICE. Our results may give evidence whether and, if so, to what extent BIs affect GBA/IQWiG EBAs and NICE STAs, and could further advance empirical HTA research. Insights gained from our article will be of relevance to HTA experts as well as to decision makers acting in the German or British healthcare sectors.

Methods

Databases

GBA/IQWiG and NICE publish official documents related to assessments and appraisals on their respective websites. First, we identified all GBA appraisals[23] and IQWiG assessments[23] completed between January 1, 2011, and June 30, 2018, as well as all NICE STAs[9] issued during the same period. We then systematically screened the relevant documents for evidence by HTA outcome referring to relevant data by patient subgroup level from GBA/IQWiG EBAs and NICE STAs.

From GBA appraisals and IQWiG assessment reports, we included data as follows (see Appendix Table 1 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2023.02.018): (1) benefit determination result by evaluation category in terms of certainty (proof, indication, hint) and extent (with added benefit: major, considerable, minor, nonquantifiable; without added benefit: no added benefit, lesser benefit) as the outcome variable and (2) multiple explanatory variables, including publication date (year); therapeutic area (14 indications); the availability of an appropriate comparative therapy (ACT) (yes/no) and study evidence in terms of relevant randomized controlled trials (RCTs) (yes/no); clinical evidence by patient-relevant endpoints (yes/no), focusing on significant differences in mortality, morbidity, or health-related quality of life; and the orphan drug (OD) status (yes/no). Evidence related to patient population size (treatable SHI target population with or without additional benefit) and annual treatment cost per patient (in €, including the additionally required SHI services) were considered for both the new drug and the ACT. We referred to a value range (minimum, maximum) and calculated the mean (average); if no value range was found, we included a single value (average) from the reference documents. Furthermore, potential discrepancies in the definition of patient subgroups by GBA and IQWiG were adjusted referring to the respective decision level in GBA appraisals. In the case that variations for assessment outcomes were identified, we considered the most favorable recommendation by IQWiG.

From NICE guidance, we included the following information and, in case of missing data, we additionally considered independent Evidence Review Group reports from the NICE website (see Appendix Table 2 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2023.02.018): (1) final recommendation (recommended/restricted/not recommended) as the outcome variable and (2) multiple explanatory variables, including publication date (year); type of condition (14 therapeutic areas); clinical effectiveness (worse/comparable/better), derived from the size of the effectiveness including strength of supporting evidence; CE (< £20 000/QALY, £20 000-30 000/QALY, or > £30 000/QALY), derived from the ICER (cost per QALY gained); the availability of relevant RCTs (yes/no) and a comparator drug (yes/no), both extracted from manufacturer submission documents; and criteria for EoL treatments (met/unmet). For both the new technology and the comparative therapy, the patient population size (eligible population for treatment in line with NICE recommendation) and annual drug acquisition cost per patient (in £) were included. If available, we included the indicated value range

Table 1. Descriptive statistics for GBA appraisals and IQWiG assessments.

Variable	GBA decision			IQWiG recommendation		
	No additional benefit, %	Additional benefit, %	Total, n	No additional benefit, %	Additional benefit, %	Total, n
Orphan drug status						
No	68.6	31.4	510	76.1	23.9	506
Yes	0.0	100.0	68	0.0	100.0	68
Appropriate comparative therapy in clinical study evidence						
Not available	8.8	91.3	80	31.1	68.9	106
Available	68.9	31.1	498	75.2	24.8	468
Clinical study evidence: randomized-controlled trial (included in a systematic review or meta-analysis)						
Not available	95.3	4.7	274	98.7	1.3	307
Available	37.7	62.3	236	41.0	59.0	200
Comparative effectiveness: significant difference in patient-relevant endpoints (mortality, morbidity, and health-related quality of life)						
Not available	98.7	1.3	301	99.7	0.3	338
Available	25.4	74.6	209	28.4	71.6	169
Publication year						
2011	75.0	25.0	4	75.0	25.0	4
2012	29.0	71.0	31	48.4	51.6	31
2013	70.6	29.4	51	69.4	30.6	49
2014	58.7	41.3	63	68.3	31.7	63
2015	57.6	42.4	125	69.4	30.6	124
2016	65.8	34.2	149	71.6	28.4	148
2017	63.5	36.5	96	66.7	33.3	96
2018 (January-June)	57.6	42.4	59	57.6	42.4	59
Indication/therapeutic area						
Blood and immune system	71.4	28.6	7	71.4	28.6	7
Cardiovascular	45.0	55.0	20	50.0	50.0	20
Digestive	62.5	37.5	8	62.5	37.5	8
Eye	72.7	27.3	11	72.7	27.3	11
Infection	55.1	44.9	118	72.0	28.0	118
Mental health	100.0	0.0	7	100.0	0.0	7
Metabolic	77.6	22.4	125	79.5	20.5	122
Musculoskeletal	86.4	13.6	22	86.4	13.6	22
Neurological	57.7	42.3	26	65.4	34.6	26
Oncological	52.3	47.7	176	56.8	43.2	176
Respiratory	50.0	50.0	26	61.5	38.5	26
Skin	29.4	70.6	17	31.3	68.8	16
Urological	66.7	33.3	3	66.7	33.3	3
Other	66.7	33.3	12	75.0	25.0	12
Total	60.6	39.4	578	67.9	32.1	574

	No additional benefit, average	Additional benefit, average	Total, average	No additional benefit, average	Additional benefit, average	Total, average
Annual treatment cost per patient in €, thousands						
Mean	354.4	134.6	267.3	48.8	140.5	79.4
Mean SD	2819.5	290.8	2199.3	74.2	296.0	186.1
Median	15.4	74.0	41.2	19.7	64.9	39.0
Treatable social health insurance target population in thousands						
Mean	141.6	68.3	112.6	140.0	82.6	120.7
Mean SD	495.2	282.9	425.5	487.9	315.1	438.1
Median	16.3	2.9	6.2	11.4	3.0	6.4
Budget impact estimation in €, millions						
Mean	1620	472	1160	543	618	568
Mean SD	12 100	911	9460	1470	1350	1430
Median	143	161	150	144	207	158
Incremental budget impact estimation in €, millions						
Mean	1340	262	922	190	317	233
Mean SD	12 100	596	9440	1120	891	1050
Median	22	77	43	23	79	41

GBA indicates Federal Joint Committee (Gemeinsamer Bundesausschuss); IQWiG, Institute for Quality and Efficiency in Health Care (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen).

(minimum, maximum) and calculated the mean (average); if not, we referred to a single value (average) in the reference document. As an additional outcome variable, we included CDF reconsiderations (ie, after its relaunch in July 2016).

Based on respective patient population and costing data, we estimated different BI-related scenarios for GBA and IQWiG as well as NICE (cf. Appendix in Supplemental Materials found at https://doi.org/10.1016/j.jval.2023.02.018): (1) 3 scenarios (minimum, average, maximum) for the potential BI (representing the absolute cost impact) of both the new and the comparator drug and (2) 1 scenario for the incremental BI (average potential BI of the new drug minus average potential BI of the comparator drug). In the case that no information was available for the comparator drug used by NICE, we followed the methodology used by Mauskopf et al[16] and multiplied the BI of the new technology with 1 (no comparator), 0.5 (comparator from different drug class), or 0.33 (comparator from same drug class). Although we included potential BI estimations from submitted manufacturer dossiers in Germany, data from submission documents in England were not fully accessible on the NICE website.

Regression Analyses

The role of the BI on HTA outcomes is assessed by means of regression analyses (using Stata Statistical Software: Release 15). In particular, the empirical strategy models the recommendations met by each agency as a function of the BI and other evaluation criteria. The analysis is performed for GBA, IQWiG, and NICE independently, whereas an additional model specification (for NICE) allows for reconsiderations via the CDF (cf. Appendix in Supplemental Materials found at https://doi.org/10.1016/j.jval.2023.02.018).

The following logistic regression equation models for GBA and IQWiG EBAs:

$$Outcome_{ij} = \beta_0 + \beta_1 * Criteria_{ij} + \beta_2 * Budget_{ij} + \mu_i + \tau_i + \varepsilon_{ij} \quad (1)$$

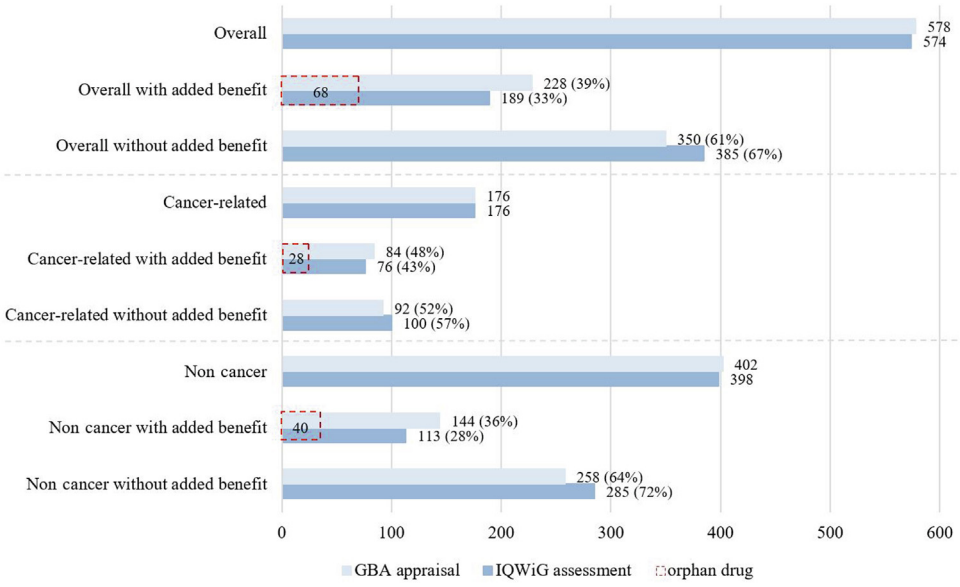
where $Outcome_{ij}$ is a 2-level categorical variable that assigns the value of 1 if either GBA or IQWiG evaluates technology i for subgroup j with added benefit and the value of 0 if they evaluate it without added benefit. $Criteria_{ij}$ includes evaluation criteria variables that are of interest for GBA or IQWiG EBAs, namely, the availability of an ACT and relevant RCTs, comparative effectiveness (patient-relevant endpoints), and the OD status. $Budget_{ij}$ is composed of the different variables addressing BI: minimum, average, and maximum estimates of the potential BI amount arising from delivering technology i to subgroup j, measured in logarithmic form, and the respective incremental BI, measured in logarithmic form. βs are the coefficients to be estimated. Year of publication and therapeutic dummy variables are denoted by μ_i and τ_i respectively, whereas ε_{ij} is the error term.

In a similar way, the ordered logistic regression question below models for NICE STAs and, in addition, for STAs including CDF reconsiderations:

$$Outcome_{ij} = \beta_0 + \beta_1 * Criteria_{ij} + \beta_2 * Budget_{ij} + \mu_i + \tau_i + \varepsilon_{ij} \quad (2)$$

where $Outcome_{ij}$ is a 3-level categorical variable that takes the value of 2 if NICE recommends technology i to subpopulation j, the value of 1 if it restrictively recommends it, and the value of 0 if it does not recommend it. In a separate specification, $Outcome_{ij}$ takes in addition the value of 2 if technology i is not recommended to subpopulation j but it is reconsidered via the CDF. Similarly, $Criteria_{ij}$ considers other assessment criteria for NICE, such as the availability of a comparator drug and relevant RCTs, the ICER (cost per QALY gained), clinical effectiveness, and the fulfillment of EoL criteria. As before, $Budget_{ij}$ takes into account the minimum, average, and maximum estimates of the potential BI from providing technology i to subpopulation j, measured in logarithmic form, and the corresponding incremental BI, measured in logarithmic form. The sample correlation between the ICER variable and the potential BI variable in the different scenarios was calculated, and its absolute value is never larger than 0.2;

Figure 1. Outcomes from GBA appraisals and IQWiG assessments by therapeutic area.



GBA indicates Federal Joint Committee (Gemeinsamer Bundesausschuss); IQWiG, Institute for Quality and Efficiency in Health Care (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen).

Table 2. Regression analyses results for GBA appraisals and IQWiG assessments.

Variable	GBA logit model								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Appropriate comparative therapy (dummy)	−0.677	−0.691	−0.774	−0.699	−0.763	−0.706	−0.756	–	–
	1.143	1.142	1.147	1.140	1.138	1.139	1.134	–	–
Clinical evidence study/RCT (dummy)	0.534	0.416	0.388	0.402	0.366	0.392	0.357	0.458	0.397
	0.561	0.609	0.612	0.607	0.611	0.606	0.611	0.608	0.627
Comparative effectiveness (dummy)	5.810*	5.845*	5.900*	5.844*	5.909*	5.844*	5.905*	5.694*	5.779*
	0.716	0.738	0.749	0.738	0.750	0.738	0.749	0.733	0.746
Orphan drug status (dummy)	–	–	–	–	–	–	–	–	–
	–	–	–	–	–	–	–	–	–
Budget impact minimum (log)		−0.001	0.483						
		0.095	0.933						
Budget impact minimum $(\log)^2$			−0.014						
			0.027						
Budget impact average (log)				0.012	0.664				0.860
				0.099	1.042				1.075
Budget impact average $(\log)^2$					−0.018				−0.024
					0.029				0.030
Budget impact maximum (log)						0.021	0.687		
						0.098	1.075		
Budget impact maximum $(\log)^2$							−0.019		
							0.030		
Increased budget impact average (log)								0.297	0.300
								0.228	0.230
Constant	−5.445	−5.443	−9.640	−5.627	−11.442	−5.767	−11.781	−6.034†	−13.812
	3.457	3.677	8.845	3.709	9.981	3.717	10.388	3.014	10.254
Publication year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Indication/therapeutic area dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	493	486	486	486	486	486	486	474	474

GBA indicates Federal Joint Committee (Gemeinsamer Bundesausschuss); IQWiG, Institute for Quality and Efficiency in Health Care (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen); RCT, randomized controlled trial.
Regression results are statistically significant at
‡ $P < .1$.
* $P < .01$.
† $P < .05$.

therefore, we can expect that the inclusion of both ICER and potential BI variables in the regression equation should not lead to multicollinearity in our model. Coefficients to be estimated are indicated by the βs, whereas year of publication and therapeutic dummy variables by μ_i and τ_i, respectively, and ε_{ij} is the error term.

Finally, robustness of estimated coefficients is tested by censoring the highest 5% BI observations in Eqs. (1) and (2).

Results

GBA Appraisals and IQWiG Assessments

Descriptive findings (Table 1) show that GBA published 262 resolutions with 578 EBAs (including 53 ODs with 68 EBAs). Because we did not identify recommendations for ivermectin (skin therapy), linagliptin and gaxilose (metabolic therapies), only 574 recommendations were considered from IQWiG assessments. Overall, we found 228 of 578 GBA decisions (40%) and 189 of 574 IQWiG recommendations (33%) with additional benefit. Cancer drugs represent the largest group by therapeutic area (99 drugs with 176 EBAs, 30%). GBA confirmed additional benefit for 84 of 176 cancer drugs (48% including 28 ODs), whereas IQWiG recommended 76 of 176 cancer therapies (43%). For noncancer drugs (163 drugs with 402 EBAs, 70%), metabolic drugs (125 EBAs) were most frequent followed by treatments for infections (118 EBAs) and neurological conditions (26 EBAs). Although GBA reported additional benefit for 144 of 402 noncancer drugs (36%, including 40 ODs), IQWiG recommended 113 of 398 noncancer therapies (28%) (Fig. 1).

By comparison, GBA decisions and IQWiG recommendations agreed for the majority of EBA outcomes (87%). This again is in line with the concordance of outcomes for patient-relevant endpoints (87%) as the main criterion for comparative effectiveness. Variations primarily exist with regard to annual treatment cost per patient (mean, GBA €267 279 vs IQWiG €79 400). GBA appraisals showed higher costs for treatments without added benefit (mean,

Table 2. Continued

IQWiG logit model								
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
−0.185	−0.266	−0.232	−0.126	−0.098	−0.094	−0.081	–	–
1.105	1.243	1.241	1.248	1.246	1.248	1.248	–	–
2.810*	2.549*	2.503*	2.738*	2.717*	2.800*	2.789*	2.935*	2.901*
0.811	0.788	0.794	0.799	0.802	0.805	0.808	0.833	0.830
6.175*	6.449*	6.467*	6.326*	6.318*	6.301*	6.296*	5.858*	6.263*
1.159	1.233	1.231	1.210	1.206	1.207	1.205	1.153	1.227
–	–	–	–	–	–	–	–	–
–	–	–	–	–	–	–	–	–
	0.254†	0.776						
	0.116	1.158						
		−0.015						
		0.032						
			0.278†	0.714				0.418
			0.129	1.377				1.392
				−0.012				−0.002
				0.037				0.038
					0.279†	0.501		
					0.130	1.419		
						−0.006		
						0.038		
							0.022	−0.167
							0.217	0.232
−6.739	−11.071*	−15.728	−12.068*	−16.091	−12.257*	−14.331	−7.678*	−14.087
5.288	3.017	10.731	3.363	13.120	3.452	13.638	2.152	13.094
Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
488	470	470	471	471	471	471	443	443

€354 445 vs €134 618 with additional benefit), whereas IQWiG assessments reported higher costs for drugs with recommendation (mean, €140 532 vs without recommendation €48 834). The average size of patient populations was similar for GBA appraisals and IQWiG assessments (mean, 112 644 vs 120 731). Accordingly, we found substantial variations in the potential (mean, GBA €1160 million vs IQWiG €568 million) and the incremental BI estimations. Compared with IQWiG assessments, we identified similar costing data (mean, €80 275) and higher figures for the patient population size (mean, 156 218) with a smaller number of subgroups per intervention in the submitted manufacturer dossiers. Unsurprisingly, potential BI estimations (mean, €469 million) reported in the dossiers were lower than those estimated for IQWiG.

ODs (not exceeding SHI expenditures of €50 million per year) dropped out of the regression model for both GBA and IQWiG as (by definition) variations in EBA outcomes were nonexistent. Results from Eq. (1) indicate that differences in patient-relevant endpoints (comparative effectiveness) had a positive significant impact on both GBA decisions ($P < .01$) and IQWiG recommendations ($P < .01$). In addition, IQWiG assessments were positively associated with the availability of relevant RCTs ($P < .01$). From dummy variables publication date and therapeutic area, years 2014 to 2017 ($P < .05$) and therapies for skin conditions ($P < .05$) turned out to be significant in IQWiG assessments, whereas treatments for oncological and hematological diseases ($P < .1$) and skin conditions ($P < .05$) were significant in GBA appraisals (Table 2).

Although the different scenarios for the potential BI did not show any significant association with GBA appraisals ($P > .1$), all 3 scenarios (minimum, average, maximum) correlated with IQWiG assessments ($P < .05$, positive linear relationship between the potential BI and recommendations). IQWiG assessments were more rigorous than GBA appraisals regarding the consideration of evidence-based evaluation criteria (ie, evaluation of submitted clinical study evidence tended to be more rigorous in drugs with higher annual treatment costs). Due to its statutory role within the German healthcare system, the GBA also considers additional factors such as stakeholder involvement—and thus, the potential

Table 3. Descriptive statistics for NICA STA recommendations.

Variable	NICE recommendation				NICE recommendation considering CDF reimbursement			
	Not recommended, %	Recommended with restriction, %	Recommended, %	Total, n	Not recommended, %	Recommended with restriction, %	Recommended, %	Total, n
Comparator drug in clinical study evidence								
Not available	60.0	30.0	10.0	20	10.0	30.0	60.0	20
Available	15.4	61.5	23.1	247	9.7	61.5	28.7	247
Clinical study evidence: randomized controlled trial (included in a systematic review or meta-analysis)								
Not available	42.9	42.9	14.3	14	7.1	42.9	50.0	14
Available	17.4	60.1	22.5	253	9.9	60.1	30.0	253
ICER: most plausible estimation or range								
< £20 000/QALY	1.3	69.7	28.9	76	1.3	69.7	28.9	76
£20 000-30 000/QALY	3.8	73.1	23.1	78	1.3	73.1	25.6	78
> £30 000/QALY	43.6	40.6	15.8	101	22.8	40.6	36.6	101
NICE assessment of clinical effectiveness based on submitted clinical study evidence								
Worse	0.0	100.0	0.0	2	0.0	100.0	0.0	2
Similar	19.6	60.8	19.6	158	13.9	60.8	25.3	158
Better	17.8	56.1	26.2	107	3.7	56.1	40.2	107
End-of-life criteria (if applicable)								
Not met	53.6	25.0	21.4	28	35.7	25.0	39.3	28
Met	33.9	45.8	20.3	59	5.1	45.8	49.2	59
Publication year								
2011	13.8	65.5	20.7	29	13.8	65.5	20.7	29
2012	5.9	58.8	35.3	17	5.9	58.8	35.3	17
2013	22.2	55.6	22.2	18	22.2	55.6	22.2	18
2014	5.3	42.1	52.6	19	5.3	42.1	52.6	19
2015	8.7	69.6	21.7	46	8.7	69.6	21.7	46
2016	26.1	50.0	23.9	46	13.0	50.0	37.0	46
2017	25.0	65.6	9.4	64	4.7	65.6	29.7	64
2018 (January-June)	28.6	50.0	21.4	28	10.7	50.0	39.3	28
Indication/therapeutic area								
Blood and immune system	14.3	71.4	14.3	7	14.3	71.4	14.3	7
Cardiovascular	0.0	52.0	48.0	25	0.0	52.0	48.0	25
Digestive	0.0	83.3	16.7	6	0.0	83.3	16.7	6
Eye	0.0	68.8	31.3	16	0.0	68.8	31.3	16
Infection	10.0	86.7	3.3	30	10.0	86.7	3.3	30
Mental health	0.0	50.0	50.0	2	0.0	50.0	50.0	2
Metabolic	12.5	75.0	12.5	16	12.5	75.0	12.5	16
Musculoskeletal	4.5	90.9	4.5	22	4.5	90.9	4.5	22
Neurological	0.0	50.0	50.0	6	0.0	50.0	50.0	6
Oncological	36.0	38.6	25.4	114	14.9	38.6	46.5	114
Respiratory	0.0	100.0	0.0	5	0.0	100.0	0.0	5
Skin	0.0	100.0	0.0	8	0.0	100.0	0.0	8
Urological	0.0	100.0	0.0	1	0.0	100.0	0.0	1
Other	22.2	44.4	33.3	9	22.2	44.4	33.3	9
Total	18.7	59.2	22.1	267	9.7	59.2	31.1	267

	Not recommended, average	Recommended with restriction, average	Recommended, average	Total, average	Not recommended, average	Recommended with restriction, average	Recommended, average	Total, average
Annual drug acquisition cost per patient in £, thousands								
Mean	43.3	26.0	28.8	29.8	31.5	26.0	36.6	29.8
Mean SD	28.4	27.8	35.8	30.5	28.1	27.8	34.9	30.5
Median	43.1	14.5	15.8	20.4	28.7	14.5	30.8	20.4
Eligible population for treatment in line with NICE recommendation in thousands								
Mean	335.0	110.8	38.6	136.8	643.7	110.8	27.6	136.8
Mean SD	2261.4	921.7	155.0	1207.7	3133.1	921.7	131.5	1207.7
Median	0.7	1.7	1.3	1.4	1.6	1.7	0.7	1.4
Budget impact estimation in £, millions								
Mean	332	125	62	149	601	125	52	149
Mean SD	2000	612	152	978	2740	612	131	978
Median	24	27	19	24	31	27	19	24
Incremental budget impact estimation in £, millions								
Mean	170	3	1	33	304	3	6	33

continued on next page

Table 3. Continued

	Not recommended, average	Recommended with restriction, average	Recommended, average	Total, average	Not recommended, average	Recommended with restriction, average	Recommended, average	Total, average
Mean SD	997	188	113	457	1370	188	97	457
Median	13	4	5	6	11	4	10	6

CDF indicates Cancer Drugs Fund; ICER, incremental cost-effectiveness ratio; NICE, National Institute for Health and Care Excellence; QALY, quality-adjusted life year.

BI seems to be of (practical) relevance to the final decisions but not to IQWiG recommendations. Interestingly, we found no significant association for neither GBA appraisals nor IQWiG assessments with the incremental BI ($P > .1$). Finally, results remained nearly the same when excluding outliers from the regressions.

NICE STA Guidance

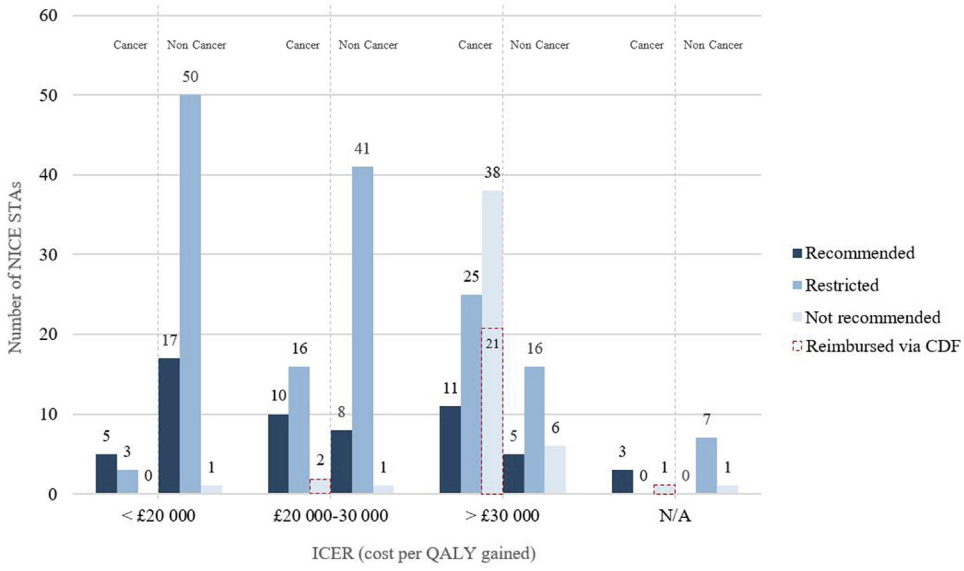
Descriptive findings (Table 3) reveal that NICE completed 207 STAs including 267 recommendations. NICE recommended 217 of 267 technologies (81%) for routine use in the NHS, of which 158 drugs (73%) were restricted. Cancer-related technologies are the largest group by indication area (114 of 267, 43%). NICE recommended 73 of 114 cancer-related technologies (64%) including 44 of 73 restricted appraisals (60%). Nevertheless, 24 of 41 cancer technologies (59%) not recommended by NICE were reconsidered by the CDF. Of 153 noncancer technologies, NICE recommended 144 STAs (94%) including 114 of 144 restrictions (79%). Most of the recommendations refer to technologies for infectious (30 STAs), cardiovascular (25 STAs), and musculoskeletal disorders (22 STAs).

Annual drug acquisition costs and the size of patient population were higher for technologies not recommended (mean, £43 306 and 334 954 patients) than those restricted (mean, £25 974 and 110 763 patients) or recommended (mean, £28 809 and 38 626 patients) by NICE for use within the NHS. This again seems to be reflected by estimations for the potential (average in millions: not recommended £332; restricted £125; recommended £62) and the incremental BI. When including CDF reconsiderations, the costing data for technologies recommended by NICE (mean, £36 634) were higher than restricted (mean, £25 974) or not recommended therapies (mean, £31 451). Nevertheless, estimations for the potential (average in millions: not recommended £601; restricted £125; recommended £52) and the incremental BI remained similar, primarily affected by patient population data.

Results from Eq. (2) reveal that NICE recommendations are largely driven by cost per QALY gained (Fig. 2). An increasing ICER (> £30 000/QALY) raised the probability for a negative recommendation significantly ($P < .01$). In addition, STA outcomes correlated with the availability of a relevant comparator drug ($P < .05$), and if applicable, the consideration of EoL criteria in cancer-related technologies was positively associated with recommendations by NICE ($P < .01$). From the dummy variables publication date and type of condition, years 2014 and 2015 ($P < .05$) and technologies for infectious, metabolic, skin, and cancer diseases ($P < .05$) turned out to be significant in NICE STAs.

Our findings indicate that 2 potential BI scenarios (average, maximum) correlated with recommendations by NICE. In fact, NICE seemed to be extremely strict in its approach to high cost (per QALY) technologies ($P < .01$, inverted-U relationship of the

Figure 2. NICE STA recommendations by (incremental) cost per QALY gained.



CDF indicates Cancer Drugs Fund; ICER, incremental cost-effectiveness ratio; N/A, not applicable; NICE, National Institute for Health and Care Excellence; QALY, quality-adjusted life-year; STA, single technology appraisal.

Table 4. Regression analyses results for NICE STA recommendations.

Variables	NICE ordered logit model							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Comparator drug (dummy)	1.469†	1.464†	1.382†	1.437†	1.403†	1.502†	1.401†	1.495†
	0.584	0.710	0.596	0.601	0.592	0.600	0.592	0.600
Clinical evidence study/RCT (dummy)	−0.285	−0.878	−0.127	−0.090	−0.224	−0.172	−0.276	−0.220
	0.690	0.799	0.692	0.693	0.697	0.700	0.699	0.702
ICER/£20 000-30 000/QALY (dummy)	−0.390	0.087	−0.468	−0.448	−0.406	−0.381	−0.370	−0.356
	0.402	1.901	0.409	0.409	0.407	0.407	0.408	0.407
ICER/> £30 000/QALY (dummy)	−2.341*	−3.166	−2.473*	−2.457*	−2.344*	−2.345*	−2.324*	−2.335*
	0.459	2.017	0.470	0.470	0.461	0.462	0.460	0.461
Clinical effectivenes/similar (dummy)	0.133	–	0.833	0.492	0.167	0.230	0.096	0.177
	2.248	–	2.315	2.347	2.253	2.267	2.251	2.264
Clinical effectiveness/better (dummy)	0.831	−0.233	1.475	1.099	0.851	0.864	0.783	0.825
	2.229	0.543	2.295	2.334	2.233	2.247	2.232	2.244
End-of-life criteria (dummy)		2.561*						
		0.729						
Budget impact minimum (log)			−0.137	0.672				
			0.088	0.863				
Budget impact minimum (log)^2				−0.024				
				0.026				
Budget impact average (log)					−0.020	1.508‡		
					0.092	0.887		
Budget impact average (log)^2						−0.046‡		
						0.026		
Budget impact maximum (log)							0.030	1.473‡
							0.093	0.874
Budget impact maximum (log)^2								−0.043‡
								0.026
Increased budget impact average (log)								
Intercept 1	−2.347	−17.652	−3.409	2.672	−2.590	9.701	−2.018	9.656
	2.414	4413.685	2.554	6.938	2.626	7.582	2.640	7.525
Intercept 2	1.276	−14.967	0.288	6.381	1.031	13.359‡	1.606	13.313‡
	2.407	4413.685	2.540	6.950	2.620	7.613	2.639	7.557
Publication year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Indication/therapeutic area dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	255	85	251	251	254	254	254	254

CDF indicates Cancer Drugs Fund; ICER, incremental cost-effectiveness ratio; NICE, National Institute for Health and Care Excellence; QALY, quality-adjusted life year; RCT, randomized controlled trial.
Regression results are statistically significant at
*$P < .01$.
†$P < .05$.
‡$P < .1$.

potential BI and STA outcomes). Nevertheless, this effect partly disappeared given that expensive cancer drugs can be reconsidered by the CDF (Table 4). Interestingly, we did not find significant correlations with the incremental BI scenario ($P > .1$). Results for the different BI estimations did not show significant differences when excluding outliers.

Regression results including CDF reconsiderations remained very similar. Nevertheless, our findings suggest that clinical effectiveness played a more significant role in STA outcomes reimbursed via the CDF and, consequently, CE might have been less crucial for positive recommendations by NICE. Apparently, the relaunch of the CDF improved HTAs of cancer-related technologies

Table 4. Continued

NICE ordered logit model		NICE ordered logit model considering CDF reimbursement									
(9)	(10)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
–	–	−1.448†	−1.829†	−1.308†	−1.254†	−1.284†	−1.218†	−1.289†	−1.228†	–	–
–	–	0.586	0.730	0.595	0.599	0.597	0.600	0.598	0.601	–	–
0.458	0.462	−0.183	−1.224	−0.233	−0.204	−0.274	−0.247	−0.292	−0.262	−0.625	−0.687
0.865	0.873	0.673	0.834	0.678	0.680	0.681	0.684	0.683	0.686	0.863	0.883
−0.385	−0.360	−0.471	0.024	−0.492	−0.478	−0.470	−0.455	−0.459	−0.450	−0.544	−0.508
0.415	0.421	0.412	1.948	0.417	0.415	0.416	0.415	0.417	0.415	0.427	0.432
−2.373*	−2.387*	−1.913*	−2.396	−1.999*	−1.982*	−1.929*	−1.926*	−1.923*	−1.927*	−2.072*	−2.070*
0.472	0.475	0.455	2.084	0.460	0.461	0.457	0.457	0.457	0.457	0.474	0.479
0.010	0.069	−0.652	–	−0.377	−0.759	−0.652	−0.594	−0.673	−0.602	−0.393	−0.379
2.307	2.333	2.300	–	2.351	2.389	2.309	2.324	2.306	2.319	2.391	2.425
0.729	0.733	0.752	1.593*	1.037	0.618	0.786	0.802	0.765	0.803	1.039	0.998
2.285	2.310	2.283	0.562	2.334	2.377	2.291	2.306	2.289	2.302	2.372	2.406
			1.815*								
			0.693								
				−0.054	0.868						
				0.090	0.893						
					−0.028						
					0.027						
	1.639‡					−0.006	1.333				1.920†
	0.914					0.095	0.901				0.951
	−0.049‡						−0.040				−0.056†
	0.027						0.027				0.028
								0.011	1.236		
								0.095	0.892		
									−0.036		
									0.026		
0.116	0.126									0.055	0.058
0.123	0.125									0.124	0.125
−3.060	9.878	−4.796‡	−17.392	−5.099‡	1.844	−4.815‡	5.953	−4.618‡	5.282	−4.265	11.048
2.551	7.706	2.483	1809.867	2.607	7.168	2.705	7.675	2.719	7.652	2.646	8.000
0.684	13.673‡	−1.038	−14.522	−1.303	5.662	−1.047	9.760	−0.850	9.083	−0.312	15.071‡
2.541	7.744	2.463	1809.867	2.585	7.190	2.686	7.709	2.702	7.685	2.627	8.055
Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
235	235	255	85	251	251	254	254	254	254	235	235

compared with outcomes from previous years (ie, from January 2011 to June 2016).

Discussion

Findings in this article indicate that both GBA/IQWiG and NICE follow their official evaluation criteria consistently. Supported from comparative analyses of matched drug pairs,[24,25] existing differences in HTA outcomes are well explained by official evaluation criteria and exceptional regulations used in the respective assessment process, as well as by institutional contexts and healthcare systems (English NHS vs German SHI). In contrast, the BI seems to play a (potential) role in HTA outcomes, although it is not (yet) an official evaluation criterion in both GBA/IQWiG EBAs and NICE STAs. In fact, we observed associations of the potential BI estimations and HTA outcomes, but not for the scenarios based on the incremental BI. This might be partly explained by variations in comparator drug data and by assumptions that have been made to complete missing information.

Our results confirm in principle the findings by Mauskopf et al[16] that the potential BI correlated (negatively) with NICE recommendations, also when controlling for CE and other predictors that have shown to be associated with STA outcomes. Nevertheless, this effect partly disappears when including the CDF variable; thus, the CDF relaunch seems to have an impact on (timely restricted) patient access toward expensive cancer drugs.[15,26]

Interestingly, the potential BI was positively associated with IQWiG recommendations, but not with GBA decisions. Although comparative effectiveness in terms of patient-relevant endpoints has shown to be the key evaluation criteria in both GBA appraisals and IQWiG assessments, IQWiG seems to be more strict toward the consideration of evidence-based evaluation criteria than GBA.[21,24,27,28] Dintsios et al[28] also indicated that GBA seems to act as a "corrective" of IQWiG recommendations. Given the fact that GBA appraisals consider additional evaluation factors as well (eg, stakeholder hearings as an official component of the AMNOG process), we assume that decisions might be implicitly influenced by the potential BI (ie, negative association of GBA appraisals and the potential BI). In addition, in contrast to Fischer and Stargardt,[21] we observed higher costing estimations for positive (compared with negative) assessments by IQWiG, but not for GBA appraisals. Overall, IQWiG recommendations and GBA decisions showed high concordance, although IQWiG may change its recommendations within commissioned addenda.[28]

In line with former empirical HTA studies,[24,25,29-33] our data set reveals differences in HTA outcomes by therapeutic area. While NICE STAs, for example, were relatively more positive toward treatments for metabolic and musculoskeletal disorders, GBA/IQWiG EBAs were more favorable toward cancer-related drugs. NICE seems to be more flexible with regard to the submission of non-RCTs or indirect comparison studies than GBA/IQWiG.[34-36] Another reason might be that NICE STAs also include modeling of long-term outcomes beyond the time horizon reported in clinical studies. Although GBA/IQWiG EBAs rely primarily on evidence-based results from RCTs, the basic requirements for submitted clinical evidence in NICE guidelines seem to be relatively similar.[35,36]

Nevertheless, cancer-related GBA appraisals and IQWiG assessments seem to be largely driven by disease morbidity and survival benefit, such as progression-free survival,[27] whereas NICE heavily relies on cost per patient by applying a cost per QALY threshold.[11-15,37] Similar to Dakin et al,[12,13] we found that CE is the major, although not the sole, driver of NICE recommendations. Given that noncancer STAs have been more favorable than cancer-related recommendations, the ICER criterion might be relaxed under EoL considerations.[38] This again is similar for STAs published after the relaunch of the CDF, because the overall approval rate of (very) expensive cancer drugs by NICE has slightly increased.[15,26,39]

Underlying value judgments of healthcare policy makers might be reflected by national HTA procedures and their exceptional regulations, such as CDF reconsiderations in England or OD designations in Germany. Thus, the contentious issue remains whether and, if so, to what extent BIAs shall be included in HTA to inform policy and decision makers.[4,40,41] Some scholars, for example, argue that drugs for rare and ultrarare diseases usually cannot meet conventional benchmarks for CE, although they are affordable and barely contribute to growth in pharmaceutical expenditure.[42,43] Bilinski et al[44] emphasized the opportunities when integrating BIAs and (some sort of) CEAs in priority setting for global health programs given that some new technologies deemed cost-effective are not affordable within given budget constraints. Indeed, a different cost perspective is needed to measure social willingness to pay, and therefore, the BI might be of relevance in healthcare policy making to ensure "affordability"[41,45]—especially for highly innovative developments such as transformative cell and gene therapies.[46]

Nevertheless, others identified a growing number of BIAs in the peer-reviewed literature, in addition to an increase in published guidelines and good practices,[4,47] albeit many of these analyses were not of high-quality standards.[48,49] In particular, van de Vooren et al[49] highlighted that, in studies funded by pharmaceutical companies, BIAs seem to be tailored to indicate short-term savings induced by innovative, high-priced medicines.

To the best of our knowledge, this is the first empirical study analyzing different scenarios for the potential BI and official evaluation criteria from GBA/IQWiG in Germany and NICE in England. Our results are supported by findings of 2 HTA comparison studies[24,25] primarily focusing on matched drug pair outcomes. Indeed, ongoing research is needed to get a better understanding of BIAs and their (potential) role in HTAs published by GBA/IQWiG and NICE and in comparison with other national HTA institutions considering the BI in decision-making processes.

The main limitation of our study is that we had to estimate the different BI scenarios based on potential costing (referring to annual treatment or drug acquisition cost per patient) and patient population data from published documents by GBA/IQWiG and NICE. This is because NICE estimates on the BI are only available for technologies that are recommended for routine use in the NHS and for patient subgroups that are included in the recommendation. In addition, BI estimations from company evidence submissions are usually not available for the public.[50] In contrast, IQWiG derives the absolute cost impact of positively evaluated therapies from pharmaceutical manufacturer dossier data, and GBA does usually not report any specific information related to the BI of a new treatment. To control the different BI estimations, we tested the study samples for official evaluation criteria that have been shown to be relevant in former empirical HTA analyses.[11-16,21,24,25,27-33] In all our analyses, we focused on HTA outcomes by patient subgroups only and not by indication. In addition to STAs published before January 2011, we excluded multiple technology appraisals (MTAs) and highly specialised technologies (HSTs) published by NICE. Finally, ODs were only included in the descriptive results because they dropped out of the regression model for both GBA and IQWiG.

Conclusions

Our results indicate that GBA/IQWiG and NICE follow their official evaluation criteria in a consistent manner. Although GBA appraisals and IQWiG assessments strictly focus on the principles of EBM, NICE STAs are largely driven by the cost per QALY gained. Interestingly, after controlling for all significant variables, the BI seems to have an effect on HTA outcomes both in Germany and in England. Therefore, BIAs might be of relevance in HTA decision-making procedures followed by GBA/IQWiG and NICE, although their (unofficial) role remains ambiguous.

Supplemental Material

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jval.2023.02.018.

Article and Author Information

Accepted for Publication: February 28, 2023

Published Online: April 17, 2023

doi: https://doi.org/10.1016/j.jval.2023.02.018

Author Affiliations: Division of Health Economics, German Cancer Research Center (DKFZ), Heidelberg, Germany (Schaefer, Hernández, Schlander); Mannheim Medical Faculty, Heidelberg University, Mannheim, Germany (Schaefer, Schlander); Institute for Innovation & Valuation in Health Care (InnoVal^{HC}), Wiesbaden, Germany (Schaefer, Schlander); Heidelberg Institute of Global Health (HIGH), Heidelberg University, Heidelberg, Germany (Baernighausen); Department of Global Health and Population, Harvard T.H. Chan School of Public Health, Boston, MA, USA (Baernighausen); Interdisciplinary Center for Health Technology Assessment and Public Health (IZPH), University of Erlangen-Nürnberg, Erlangen, Germany (Kolominsky-Rabas); Alfred-Weber-Institute, Heidelberg University, Heidelberg, Germany (Schlander).

Correspondence: Ramon Schaefer, MA, BSc, German Cancer Research Center, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany. Email: ramon.schaefer@dkfz.de

Author Contributions: *Concept and design:* Schaefer, Hernández, Bärnighausen, Kolominsky-Rabas, Schlander
Data acquisition: Schaefer
Data analysis and interpretation: Schaefer, Hernández, Bärnighausen, Schlander
Manuscript draft: Schaefer, Hernández
Manuscript revision: Hernández, Bärnighausen, Kolominsky-Rabas, Schlander

Conflict of Interest Disclosures: All authors approved the final manuscript version to be published. The authors reported no conflicts of interest.

Funding/Support: The authors received no financial support for this research.

Acknowledgment: The authors thank Lorenz Selberg for his support with data acquisition and Rachel Eckford for proofreading the manuscript.

REFERENCES

1. Banta D. The development of health technology assessment. *Health Policy*. 2003;63(2):121–132.
2. Battista RN, Hodge MJ. The evolving paradigm of health technology assessment: reflections for the millennium. *CMAJ*. 1999;160(10):1464–1467.
3. Drummond MF, Schwartz JS, Jönsson B, et al. Key principles for the improved conduct of health technology assessments for resource allocation decisions. *Int J Technol Assess Health Care*. 2008;24(3):244–258.
4. Foroutan N, Tarride JE, Xie F, Levine M. A methodological review of national and transnational pharmaceutical budget impact analysis guidelines for new drug submissions. *Clinicoecon Outcomes Res*. 2018;10:821–854.
5. Mauskopf JA, Sullivan SD, Annemans L, et al. Principles of good practice for budget impact analysis: report of the ISPOR Task Force on Good Research Practices–Budget Impact Analysis. *Value Health*. 2007;10(5):336–347.
6. Sullivan SD, Mauskopf JA, Augustovski F, et al. Budget impact analysis–principles of good practice: report of the ISPOR 2012 Budget Impact Analysis Good Practice II Task Force. *Value Health*. 2014;17(1):5–14.
7. Verfahrensordnung des Gemeinsamen Bundesausschusses. Gemeinsamer Bundesausschuss. https://www.g-ba.de/richtlinien/42/. Accessed December 30, 2022.
8. Guide to the processes of technology appraisal: process and methods. National Institute for Health and Care Excellence. https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-technology-appraisal-guidance. Accessed December 12, 2021.
9. Guidance and advice list. National Institute for Health and Care Excellence. http://www.nice.org.uk/guidance/published?type=cg. Accessed January 26, 2019.
10. Charlton V. NICE and fair? Health technology assessment policy under the UK's National Institute for Health and Care Excellence, 1999-2018. *Health Care Anal*. 2020;28(3):193–227.
11. Cerri KH, Knapp M, Fernandez JL. Decision making by NICE: examining the influences of evidence, process and context. *Health Econ Policy Law*. 2014;9(2):119–141.
12. Dakin HA, Devlin NJ, Odeyemi IA. "Yes", "No" or "Yes, but"? Multinomial modelling of NICE decision-making. *Health Policy*. 2006;77(3):352–367.
13. 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.
14. Devlin N, Parkin D. Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. *Health Econ*. 2004;13(5):437–452.
15. Walton MJ, O'Connor J, Carroll C, Claxton L, Hodgson R. A review of issues affecting the efficiency of decision making in the NICE single technology appraisal process. *Pharmacoecon Open*. 2019;3(3):403–410.
16. Mauskopf J, Chirila C, Birt J, Boye KS, Bowman L. Drug reimbursement recommendations by the National Institute for Health and Clinical Excellence: have they impacted the National Health Service budget? *Health Policy*. 2013;110(1):49–59.
17. de Folter J, Trusheim M, Jonsson P, Garner S. Decision-components of NICE's technology appraisals assessment framework. *Int J Technol Assess Health Care*. 2018;34(2):163–171.
18. Shah KK, Tsuchiya A, Wailoo AJ. Valuing health at the end of life: a stated preference discrete choice experiment. *Soc Sci Med*. 2015;124:48–56.
19. Cancer Drugs Fund. NHS England. https://www.england.nhs.uk/cancer/cdf/. Accessed January 15, 2022.
20. Sozialgesetzbuch (SGB) Fünftes Buch (V) - Gesetzliche Krankenversicherung - § 35b Kosten-Nutzen-Bewertung von Arzneimitteln. Bundesministerium der Justiz. http://www.gesetze-im-internet.de/sgb_5/__35b.html. Accessed November 20, 2021.
21. Fischer K, Stargardt T. Early benefit assessment of pharmaceuticals in Germany: manufacturers' expectations versus the federal joint committee's decisions. *Med Decis Making*. 2014;34(8):1030–1047.
22. Methodenpapier: Allgemeine Methoden. IQWiG. https://www.iqwig.de/ueber-uns/methoden/methodenpapier/. Accessed December 30, 2022.
23. Nutzenbewertung von Arzneimitteln: Verfahren nach § 35a SGB V. Gemeinsamer Bundesausschuss. https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/. Accessed January 15, 2019.
24. Schaefer R, Schlander M. Is the National Institute for Health and Care Excellence (NICE) in England more 'innovation-friendly' than the Federal Joint Committee (G-BA) in Germany? *Expert Rev Pharmacoecon Outcomes Res*. 2019;19(4):453–462.
25. Schaefer R, Hernandez D, Selberg L, Schlander M. HTA in England, France and Germany: what do matched drug pairs tell us about recommendations by national HTA agencies? *J Comp Eff Res*. 2021;10(16):1187–1195.
26. Sabry-Grant C, Malottki K, Diamantopoulos A. The Cancer Drugs Fund in practice and under the new framework. *Pharmacoeconomics*. 2019;37(7):953–962.
27. Ruof J, Schwartz FW, Schulenburg JM, Dintsios CM. Early benefit assessment (EBA) in Germany: analysing decisions 18 months after introducing the new AMNOG legislation. *Eur J Health Econ*. 2014;15(6):577–589.
28. Dintsios CM, Worm F, Ruof J, Herpers M. Different interpretation of additional evidence for HTA by the commissioned HTA body and the commissioning decision maker in Germany: whenever IQWiG and Federal Joint Committee disagree. *Health Econ Rev*. 2019;9(1):35.
29. Cerri KH, Knapp M, Fernandez JL. Untangling the complexity of funding recommendations: a comparative analysis of health technology assessment outcomes in four European countries. *Pharm Med*. 2015;29(6):341–359.
30. Clement FM, Harris A, Li JJ, Yong K, Lee KM, Manns BJ. Using effectiveness and cost-effectiveness to make drug coverage decisions: a comparison of Britain, Australia, and Canada. *JAMA*. 2009;302(13):1437–1443.
31. Fischer KE, Heisser T, Stargardt T. Health benefit assessment of pharmaceuticals: an international comparison of decisions from Germany, England, Scotland and Australia. *Health Policy*. 2016;120(10):1115–1122.
32. Nicod E, Kanavos P. Commonalities and differences in HTA outcomes: a comparative analysis of five countries and implications for coverage decisions. *Health Policy*. 2012;108(2-3):167–177.
33. Nicod E, Maynou L, Visintin E, Cairns J. Why do health technology assessment drug reimbursement recommendations differ between countries? A parallel convergent mixed methods study. *Health Econ Policy Law*. 2020;15(3):386–402.
34. Lebioda A, Gasche D, Dippel FW, Theobald K, Plantör S. Relevance of indirect comparisons in the German early benefit assessment and in comparison to HTA processes in England, France and Scotland. *Health Econ Rev*. 2014;4(1):31.
35. Akehurst RL, Abadie E, Renaudin N, Sarkozy F. Variation in health technology assessment and reimbursement processes in Europe. *Value Health*. 2017;20(1):67–76.
36. Angelis A, Lange A, Kanavos P. Using health technology assessment to assess the value of new medicines: resultsof a systematic review and expert consultation across eight European countries. *Eur J Health Econ*. 2018;19(1):123–152.
37. Charlton V. The normative grounds for NICE decision-making: a narrative cross-disciplinary review of empirical studies. *Health Econ Policy Law*. 2022;17(4):444–470.
38. Pujolras LM, Cairns J. Why do some countries approve a cancer drug and others don't? *J Cancer Policy*. 2015;4(1):21–25.
39. Aggarwal A, Fojo T, Chamberlain C, Davis C, Sullivan R. Do patient access schemes for high-cost cancer drugs deliver value to society?—lessons from the NHS Cancer Drugs Fund. *Ann Oncol*. 2017;28(1):1738–1750.
40. Culyer AJ. Ethics, priorities and cancer. *J Cancer Policy*. 2017;11:6–11.

41. Caro JJ, Brazier JE, Karnon J, et al. Determining value in health technology assessment: stay the course or tack away? *Pharmacoeconomics*. 2019;37(3):293–299.
42. Schey C, Milanova T, Hutchings A. Estimating the budget impact of orphan medicines in Europe: 2010-2020. *Orphanet J Rare Dis*. 2011;6:62.
43. Schlander M, Dintsios CM, Gandjour A. Budgetary impact and cost drivers of drugs for rare and ultrarare diseases. *Value Health*. 2018;21(5):525–531.
44. Bilinski A, Neumann P, Cohen J, Thorat T, McDaniel K, Salomon JA. When cost-effective interventions are unaffordable: integrating cost-effectiveness and budget impact in priority setting for global health programs. *PLoS Med*. 2017;14(10):e1002397.
45. Richardson J, Schlander M. Health technology assessment (HTA) and economic evaluation: efficiency or fairness first. *J Mark Access Health Policy*. 2018;7(1):1557981.
46. Tunis S, Hanna E, Neumann PJ, et al. Variation in market access decisions for cell and gene therapies across the United States, Canada, and Europe. *Health Policy*. 2021;125(12):1550–1556.
47. Orlewska E, Gulácsi L. Budget-impact analyses: a critical review of published studies. *Pharmacoeconomics*. 2009;27(10):807–827.
48. Chugh Y, De Francesco M, Prinja S. Systematic literature review of guidelines on budget impact analysis for health technology assessment. *Appl Health Econ Health Policy*. 2021;19(6):825–838.
49. van de Vooren K, Duranti S, Curto A, Garattini L. A critical systematic review of budget impact analyses on drugs in the EU countries. *Appl Health Econ Health Policy*. 2014;12(1):33–40.
50. Schlander M. The use of cost-effectiveness by the National Institute for Health and Clinical Excellence (NICE): no(t yet an) exemplar of a deliberative process. *J Med Ethics*. 2008;34(7):534–539.