British Journal of Clinical Pharmacology

Br J Clin Pharmacol (2018) **84** 1146–1155 1146

ORIGINAL ARTICLE

Balancing early access with uncertainties in evidence for drugs authorized by prospective case series – systematic review of reimbursement decisions

Correspondence Susanna Wallerstedt, Department of Pharmacology, Sahlgrenska Academy, University of Gothenburg, Box 431, SE-405 30 Gothenburg, Sweden. Tel.: +46 31 786 3410; Fax: +46 31 786 3164; E-mail: susanna.wallerstedt@pharm.gu.se

Received 13 October 2017; **Revised** 18 January 2018; **Accepted** 24 January 2018

Susanna M. Wallerstedt[1,2] and Martin Henriksson[3]

[1]Department of Pharmacology, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden, [2]Department of Clinical Pharmacology, Sahlgrenska University Hospital, Gothenburg, Sweden, and [3]Department of Medical and Health Sciences, Linköping University, Linköping, Sweden

Keywords cost-effectiveness, decision making, uncertainty, evidence-based health care, prospective case series

AIMS

To review clinical and cost-effectiveness evidence underlying reimbursement decisions relating to drugs whose authorization mainly is based on evidence from prospective case series.

METHODS

A systematic review of all new drugs evaluated in 2011–2016 within a health care profession-driven resource prioritization process, with a market approval based on prospective case series, and a reimbursement decision by the Swedish Dental and Pharmaceutical Benefits Agency (TLV). Public assessment reports from the European Medicines Agency, published pivotal studies, and TLV, Scottish Medicines Consortium and National Institute of Health and Care Excellence decisions and guidance documents were reviewed.

RESULTS

Six drug cases were assessed (brentuximab vedotin, bosutinib, ponatinib, idelalisib, vismodegib, ceritinib). The validity of the pivotal studies was hampered by the use of surrogate primary outcomes and the absence of recruitment information. To quantify drug treatment effect sizes, the reimbursement agencies primarily used data from another source in indirect comparisons. TLV granted reimbursement in five cases, compared with five in five cases for Scottish Medicines Consortium and four in five cases for National Institute of Health and Care Excellence. Decision modifiers, contributing to granted reimbursement despite hugely uncertain cost-effectiveness ratios, were, for example, small population size, occasionally linked to budget impact, severity of disease, end of life and improved life expectancy.

CONCLUSION

For drugs whose authorization is based on prospective case series, most applications for reimbursement within public health care are granted. The underlying evidence has limitations over and above the design *per se*, and decision modifiers are frequently referred to in the value-based pricing decision making.

DOI:10.1111/bcp.13531

WHAT IS ALREADY KNOWN ABOUT THIS SUBJECT

- Drug treatment effect size is a crucial component in determining the health gain and overall value of a new pharmacological therapy.
- The occasional approval by the European Medicines Agency (EMA) of new drugs where the underlying evidence is based solely on prospective case series means that drugs with unknown efficacy are increasingly available for routine use in health care.
- EMA approval of drugs with unknown efficacy is challenging for national reimbursement agencies, and it is not clear how early access should balance uncertainties in evidence when deciding whether a new drug should be reimbursed under publicly funded health care.

WHAT THIS STUDY ADDS

- In six current cases where EMA drug approval was mainly based on prospective case series, several issues with study design, over and above the lack of a control group, were identified, hampering the internal and external validity of the clinical results.
- To estimate effect sizes in the absence of a comparative study, the reimbursement agencies used data from another source in (often naïve) indirect comparisons and/or compared the results reported in the single-arm pivotal study with prior treatment or nonresponders in the same study.
- Reimbursement was granted in almost all cases as hugely uncertain and often high cost-effectiveness ratios were balanced with decision modifiers such as small population size, high disease severity and end of life, suggesting that improvements in pivotal study design and clarification of the role of decision modifiers may be warranted.

Introduction

One or more *pivotal* studies constitute the main scientific evidence in the benefit–risk assessment underlying drug marketing approval in Europe by the European Medicines Agency (EMA). Often, these studies are randomized and controlled, providing relative efficacy estimates with high validity. However, to facilitate rapid access for new innovations, particularly when unmet medical needs are addressed, the scientific evidence in some EMA drug approvals is restricted to prospective case series, that is, single-arm studies without a control group [1, 2].

Although the enabling of rapid access for selected new drugs, often for orphan indications, through so-called *conditional approval* and *adaptive pathways*, may be commendable, this process is also associated with several challenges [3]. For example, data on efficacy and safety are particularly limited when drugs are approved based on single-arm studies, which is the case in about one third of the EMA conditional approvals, and half of those within oncology [4]. With such an evidence base, it is challenging for national reimbursement agencies using value-based pricing to decide whether the new drug should be made available, and at what price, within publicly funded health care. Indeed, to consider whether the price of a new drug is reasonable, health gains need to be quantified. This is a challenging task under normal circumstances, and even more so in the absence of comparative studies.

Compiling and analysing the evidence base, as well as the decisions made by the EMA and reimbursement agencies, not only provides a means to elucidate the challenges outlined above through real case studies, but also makes it possible to initiate a discussion on how to handle these challenges in the future. To our knowledge, such a review and analysis is currently lacking in the literature. Therefore, the aim of this study was to review current cases of new drugs with market authorization based on prospective case series where health care professionals have explicitly expressed that they want to use the new therapy. The focus of the study was on clinical and cost-effectiveness evidence underlying reimbursement decisions by the Swedish Dental and Pharmaceutical Benefits Agency (TLV), the Scottish Medicines Consortium (SMC), and the British National Institute of Health and Care Excellence (NICE).

Methods

We identified all cases with the following inclusion criteria: nominated for use in public health care by professional expert groups in Region Västra Götaland (the second largest region in Sweden encompassing 1.7 million inhabitants) in 2011–2016 and evaluated within a regional process where the severity of disease, the benefit–risk balance, and the level of evidence are assessed as the basis for funding prioritizations [1]; an EMA market approval based on prospective case series; and a TLV reimbursement decision available.

To elucidate the challenges of decision making regarding the identified cases, we compiled evidence underlying the EMA approval as well as rationales for reimbursement decisions. The review was based on publicly available documentation: decisions and public assessment reports (EPARs) from the EMA, published pivotal studies, and decisions and advice/guidance documents published by TLV, SMC and NICE.

Data extraction

The first author (S.M.W.) extracted data from the EMA decisions, the EPARs, and pivotal studies, and the other author (M.H.) checked these. Data extraction included indication for use, type of approval, and benefit–risk balance as described in the EPAR. Further, information was retrieved from the pivotal studies, including number of exposed patients and results regarding the primary outcome, severe as well as life-threatening adverse events and deaths, and health-

related quality of life (HRQL). The recording of HRQL data in the pivotal studies was also checked in ClinicalTrials.gov. For subjective primary outcomes, κ was retrieved, or, if not provided, backwards calculated where possible based on information provided in the publication. The studies were also assessed according to a quality checklist for prospective case series used by the Regional Health Technology Assessment (HTA) Centre in Region Västra Götaland [5].

The second author (M.H.) extracted data from the reimbursement decision by the TLV including any publicly available background documentation, as well as the SMC and NICE guidance documentation, and the other author (S.M.W.) checked these. The focus in the extractions was the health gain in terms of quality-adjusted life years (QALYs) as well as estimated drug treatment effect sizes, calculations of cost-effectiveness, and decision modifiers.

Descriptive analyses were performed using SPSS, version 20.0 (IBM SPSS Statistics for Windows, Armonk, NY, USA).

Nomenclature of targets and ligands

Key protein targets and ligands in this article are hyperlinked to corresponding entries in http://www.guidetopharmacology.org, the common portal for data from the IUPHAR/BPS Guide to PHARMACOLOGY [6], and are permanently archived in the Concise Guide to PHARMACOLOGY 2017/18.

Results

Six current drug cases were included in the review (Figure 1). They were all intended for cancer treatment, four within haematology and two within oncology (Table 1). Three drugs were classified as orphan drugs, and four EMA approvals were initially conditioned. According to the EPAR, there was consensus that the benefit–risk balance was favourable for five drugs (**brentuximab vedotin**, **bosutinib**, **ponatinib**, **idelalisib** and **ceritinib**). For the sixth drug (**vismodegib**), the majority considered this balance acceptable.

The pivotal studies are described in Table 2. For brentuximab vedotin and bosutinib, two pivotal studies were identified. For the remaining four drugs, approval was mainly based on one study. Between 58 and 444 patients were included in the studies, the median age ranging from 31 to 64 years. For four drugs, the primary outcome in the pivotal study was the objective/observed response rate according to predefined criteria, including complete and partial response, but not stable disease. For the remaining two drugs, haematological responses according to laboratory assessments were used. In six out of eight studies, deaths were systematically recorded and described. In these studies, 48 deaths were described, six of which were assessed as treatment-related and 11 of which were due to disease progression.

Internal validity problems identified in the pivotal studies included the use of surrogate primary outcomes. The reported or estimated κ value, reflecting the inter-rater agreement on the primary outcome assessment, varied between 0.23 [7] and 0.68 [8, 9]. No study had overall survival as primary outcome, and the 12-month survival varied between 29% (in a subgroup of studied patients) [10] and 97% [11]. Regarding the external validity, the recruitment of patients was not described in any of the pivotal studies.

Table 3 presents information regarding drug treatment effect sizes and cost-effectiveness in the TLV, SMC and NICE decisions and guidance documentation. Four drugs were granted reimbursement by all agencies (brentuximab vedotin, bosutinib, ponatinib, ceritinib). Rejection occurred in one case in two agencies (vismodegib: TLV, NICE), the remainder were not applied for (vismodegib: SMC; idelalisib: NICE). Decisions were conditioned by commercial in confidence discounts in one case by the TLV, in two by the SMC and in all by the NICE.

Three main approaches to quantify the drug treatment effect size were identified: (i) using data from an arm in another study or from register data as comparator, either in a naïve way or with some matching; (ii) patients acting as their own controls by comparing the results of the new drug with results achieved by prior treatment according to retrospective extraction of information in medical records; and (iii) comparing results from the single-arm pivotal study with the results of nonresponders in the same study. The first approach (indirect comparisons) was used in four out of six TLV cases,

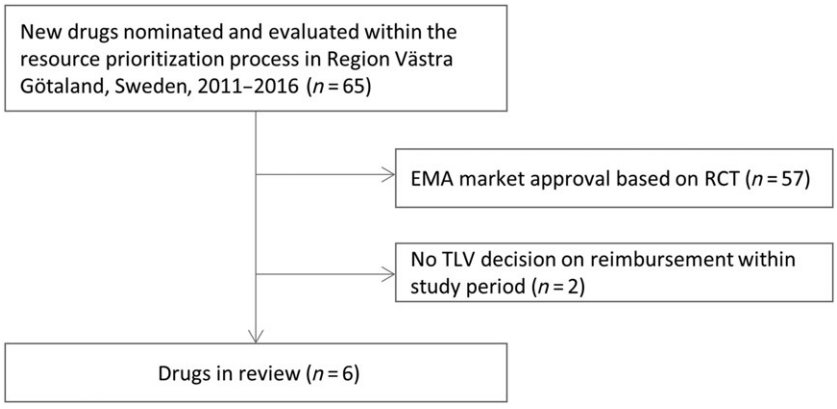

Figure 1

Flowchart of drugs included in this review, for which the European Medicines Agency (EMA) approval was based on prospective case series. RCT, randomized controlled trial; TLV, Swedish Dental and Pharmaceutical Benefits Agency

Table 1

Characteristics of studies included in the review, as well as resulting assessments by the European Medicines Agency

Substance	Product	Indication	Conditional approval	Orphan drug	Benefit/risk as described in the EPAR
Brentuximab vedotin	Adcetris	HL, AL	Y	Y	Established antitumour activity. Clinical benefit demonstrated. Acceptable risks.
Bosutinib	Bosulif	CML	Y	Y	Clinically significant benefit. Acceptable toxicity profile.
Vismodegib	Erivedge	mBCC, laBCC	2013: Y 2016: N	N	Proven antitumour activity. Established clinical benefit. Manageable toxicity.
Ponatinib	Iclusig	CML, ALL	N	Y	Very clinically relevant response rates. Manageable risks.
Idelalisib	Zydelig	FL	N	N	Convincing and clinically relevant results. At least as clinically significant as other available options. Manageable toxicity.
Ceritinib	Zykadia	NSCLC	Y	N	More efficacious compared with currently available therapies. Not well tolerated but with manageable toxicity.

AL, anaplastic lymphoma; ALL, acute lymphatic leukaemia; CML, chronic myeloid leukaemia; EPAR, European public assessment report; FL, follicular lymphoma; HL, Hodgkin lymphoma; laBCC, locally advanced basal cell cancer; mBCC, metastatic basal cell cancer; N, no; NSCLC, non-small cell lung cancer; Y, yes.

four out of five SMC cases, and in four out of five NICE cases. The second approach (prior treatment) was used in one case by all three agencies (brentuximab vedotin). The third approach (nonresponders) was used in three TLV cases (vismodegib, idelalisib, ceritinib), in none of the SMC cases, and in one NICE case (vismodegib). The uncertainties in treatment effect estimates were recognized and commented on by all agencies.

The magnitudes of estimated drug treatment effect sizes were explicitly stated in two TLV decisions and in two SMC advice and two NICE guidance documents. The remaining did not explicitly mention the patient benefit, for example concerning the estimated number of gained progression-free months. Regarding estimated health gain for the cost-effectiveness analyses, none of the pivotal studies contained information on HRQL (Table 3). According to ClinicalTrials.gov, HRQL was recorded in two pivotal studies (vismodegib, idelalisib), both with other tools than EuroQol-5 dimensions (EQ-5D) commonly used to estimate QALYs. The estimated number of QALYs gained was explicitly stated in none of the TLV decisions (although included in one publicly available background document), in four SMC advice documents and in one NICE guidance document.

The most frequent decision modifier identified in the reviewed cases, contributing to positive reimbursement decisions despite huge uncertainties in effect size estimates and cost-effectiveness ratios, was small population size, referred to in two TLV, five SMC and one NICE cases (Table 3). This modifier was accompanied by a statement on limited budget impact in one TLV and one NICE case. Other frequent decision modifiers were severity of disease, as commented upon in four granted TLV decisions, and the drug in question being classified as intended for treatment at the end of life, in three SMC and four NICE documents. Further, in three SMC cases and in one NICE case, the new drug was considered to contribute substantially to improved life expectancy.

Discussion

This review of six current cases elucidates challenges regarding reimbursement decisions for new drugs approved on the basis of single-arm studies. Indeed, regarding the underlying pivotal studies, we identified several issues over and above the lack of a control group, hampering the internal and the external validity, including the use of surrogate and at least partly subjective outcomes, and omitted information on the process of patient recruitment. Further, we found that three European reimbursement agencies using value-based pricing relied on three main approaches to quantify drug treatment effect size in the absence of a comparative study: indirect comparison using results in another clinical trial or register data; comparison with retrospectively obtained information on results of prior treatment; and comparison with a subset of nonresponders in the pivotal study. Despite huge uncertainties in treatment effect size and cost-effectiveness, reimbursement was granted in all but one TLV/NICE case, and in all available SMC cases. Important decision modifiers taken into account were, for example: small population size, occasionally linked to budget impact considerations; being intended for treatment of diseases with high severity or at end of life; and/or contributing substantially to improved life expectancy.

As surrogate primary outcomes were used in the pivotal studies in this review, and data on overall survival were not available at the time of the reimbursement decision, it may be worth emphasizing that, as far as we are aware, the surrogate outcomes used have not been correlated with overall survival for the approved indications. Indeed, most surrogate outcomes used in oncology trials have low correlation with survival [12]. Response rate, the most frequently used primary outcome in this review, has been shown to be moderately associated with overall survival in breast cancer; 63% and 71%, respectively, of the variances in two reviews were explained by other factors [13, 14]. For progression-free

Table 2

Description of pivotal studies for drugs included in the review

Substance	Pivotal study/acronym	Phase	Patients/ age*	Efficacy/survival/HRQL	AE/deaths
Brentuximab vedotin	Younes *et al.*, J Clin Oncol 2012 [8] (HL)	2	*n* = 102 31 (15–77)	PO: ORR by IRC: 75% (IRA: 0.68) OS: 22.4 months (95% CI: 21.7; NE) 12-month survival: 89% HRQL not recorded	55% AE grade ≥ 3 No deaths on treatment or within 30 days of study drug discontinuation
	Pro *et al.*, J Clin Oncol 2012 [26] (AL)	2	*n* = 58 52 (14–76)	PO: ORR by IRC: 86% (IRA: 0.65) OS: NR (95% CI: 14.6; NE) 12-month survival: 70% HRQL not recorded	60% AE grade ≥ 3 Six deaths on treatment or within 30 days of study drug discontinuation, none assessed as related to treatment: progression (*n* = 4), MI/renal failure (*n* = 1), tracheal prosthesis obstruction (*n* = 1)
Bosutinib	Cortes *et al.*, Blood 2011 [11] (second- line)	1/2	*n* = 288 53 (18–91)	PO: MCyR: 31% (IRA: NA) 12-month survival: 97% HRQL not recorded	AE not summated Deaths not described
	Khoury *et al.*, Blood 2012 [27] (third-line)	1/2	*n* = 118 56 (20–79)	PO: MCyR: 32% (IRA: NA) 12-month survival: 91% HRQL not recorded	22% AE grade 3/4 Six deaths (5%) on treatment or within 30 days of study drug discontinuation. One death assessed as related to treatment (gastrointestinal bleeding). Others: MI (*n* = 2), disease progression (*n* = 3)
Vismodegib	Sekulic *et al.*, New Engl J Med, 2012 [7] ERIVANCE	2	*n* = 96 62 (21–101)	PO: ORR by IRC: mBCC: 30% (IRA: 0.56) laBCC: 43% (IRA: 0.23) 12-month survival, final analysis: [28] 79% (mBCC), 93% (laBCC) HRQL recorded with SF-36	25% had serious AE, seven of which led to deaths assessed by the investigator as not related to treatment [ischaemic stroke (*n* = 1), MI (*n* = 1), meningeal disorder (*n* = 1), hypovolaemic shock (*n* = 1), unknown (*n* = 3)]
Ponatinib	Cortes *et al.*, New Engl J Med, 2013 [10] PACE	2	*n* = 444 59 (18–94)	CML, chronic phase; PO: MCyR: 56% (IRA: NA) CML, accelerated or blast phase, ALL; PO: MHR: 55%/31%/41% (IRA: NA) OS: NR/NR/7 months/8 months 12-month survival: 94%/ 84%/29%/40% HRQL not recorded	AE grade 3/4 not summated 18 deaths, five of which were assessed as probably/possibly related to treatment: pneumonia (*n* = 2), MI (*n* = 1), gastric haemorrhage (*n* = 1), cardiac arrest (*n* = 1). Reasons for death, not assessed as probably/possibly related to treatment, included sepsis (*n* = 4), cardiac arrest (*n* = 2), congestive heart failure (*n* = 2), cardiopulmonary failure (*n* = 1), dehydration (*n* = 1), hyperviscosity syndrome (*n* = 1), small intestine obstruction (*n* = 1) and neoplasm progression (*n* = 1)
Idelalisib	Gopal *et al.*, New Engl J Med, 2014 [9] DELTA	2	*n* = 72 64 (33–87)	PO: ORR by IRC: 54% (IRA: 0.68) 12-month survival: 80% (all types of indolent lymphoma included in study, *n* = 125) HRQL recorded with FACT-LymS	AE grade 3/4 not summated Eleven deaths on treatment or within 30 days of study drug discontinuation: disease progression (*n* = 3), pneumonia (*n* = 3), cardiac arrest (*n* = 1), cardiac failure (*n* = 1), splenic infarction (*n* = 1), sepsis (*n* = 1), and pneumonitis (*n* = 1). Treatment relation not described
Ceritinib	Shaw *et al.*, New Engl J Med, 2014 [29] ASCEND-1	1	*n* = 122 53 (22–80)	PO: MTD (IRA: NA; one assessor only) ORR by investigators: 58% 12-month survival: 65% HRQL not recorded	AE grade 3/4 not summated Deaths not described; no deaths assessed as treatment-related. In updated results, two deaths related to treatment are described: interstitial lung disease (*n* = 1), multi-organ failure in the context of infection and ischaemic hepatitis (*n* = 1) [30]. Deaths not systematically recorded

AE, adverse event; AL, anaplastic lymphoma; ALL, acute lymphatic leukaemia; CI, confidence interval; CML, chronic myeloid leukaemia; FACT-LymS, Functional Assessment of Cancer Therapy - Lymphoma; HL, Hodgkin lymphoma; HRQL, health-related quality of life; IRA, inter-rater agreement as measured by κ statistics; IRC, independent review committee; laBCC, locally advanced basal cell cancer; mBCC, metastatic basal cell cancer; MCyR, major cytogenetic response; MHR, major hematologic response; MI, myocardial infarction; MTD, maximum tolerated dose; NA, not applicable; NE, not estimated; NHR, major haematologic response; NR, not reached; ORR, overall/objective response rate; OS, overall survival; PO, primary outcome; SF-36, 36-Item short form health survey.
*median (range).

Table 3

Decisions, as well as underlying assessments of drug treatment effect sizes and cost-effectiveness, from three European national reimbursement agencies

		TLV	SMC	NICE
Brentuximab vedotin	Decision	Reimbursed with limitation	Accepted for restricted use	Recommended in the commercial access agreement
	Treatment effect	*Vs.* prior treatment; HL, HL before SCT: PFS benefit 3.7 months AL: PFS benefit 8.4 months	*Vs.* prior treatment; HL: PFS benefit: 3.7 months Indirect comparison; HL before SCT; PFS benefit not specified AL not mentioned	*Vs.* prior treatment; HL: PFS benefit: 3.7 months Indirect comparison; HL: PFS not specified AL not mentioned
	Cost-effectiveness	HL: 470 000–1 140 000 SEK/QALY; QALY benefit not specified HL before SCT: 900 000–1 450 000 SEK/QALY; QALY benefit not specified AL: assumption of lower ICER compared with HL	HL: £43 000/QALY; QALY benefit: 1.41 HL before SCT: dominating; £7000 cost saving; QALY benefit: 0.68	HL; *vs.* prior treatment: <£30 000/QALY; indirect comparison: £40 000/QALY; QALY benefit not specified
	Decision aspects	Decision modifier: High severity of disease. Conditioned: Updated health economics model required when pivotal trials are completed.	Decision modifiers: End of life, substantial improvement in life expectancy, substantial improvement in quality of life, potential to bridge to a definitive therapy, absence of other treatments, small population size.	Decision modifier: Cancer Drugs Fund applicable for a subset of patients.
Bosutinib	Decision	Reimbursed	2013: Not recommended for use 2015: Accepted for use within PAS	Recommended within PAS
	Treatment effect	Indirect comparison; equal effect	Indirect comparison; PFS benefit not specified	Indirect comparison; PFS benefit not specified
	Cost-effectiveness	Cost saving under the assumption of equal effects	Chronic phase: £39 000 or £46 000/QALY; QALY benefit: 2.05/3.79 Accelerated phase: £62 000/QALY; QALY benefit: 2.01 Blast phase: £61 000/QALY; QALY benefit: 0.85	Chronic phase: £43 000/QALY; QALY benefit not specified Accelerated phase: £58 000/QALY; QALY benefit not specified Blast phase: £60 000/QALY; QALY benefit not specified
	Decision aspects	Decision modifier: High severity of disease. Conditioned: Updated health economic analysis required when pivotal trials are completed.	Decision modifier: Small population size.	Decision modifier: End of life in accelerated and blast phases of CML.
Vismodegib	Decision	Not reimbursed	Not applied	Not recommended for use
	Treatment effect	2014: Indirect comparison; no survival benefit 2016: *vs.* nonresponders in pivotal study; no survival benefit		*Vs.* nonresponders, no survival benefit
	Cost-effectiveness	2014: 2 400 000 SEK/QALY 2016: SEK/QALY not estimated owing to uncertainty		£96 548/QALY assuming a survival benefit; QALY benefit not specified £4 694 943/QALY assuming no survival benefit; QALY benefit not specified
	Decision aspects	N/A		N/A
Ponatinib	Decision	Reimbursed	Accepted for use	Recommended within PAS
	Treatment effect	Indirect comparison; PFS benefit not specified	Indirect comparison; PFS benefit only specified for CML accelerated phase: 11.4 months	Indirect comparison (matched for CML chronic phase, naïve for other phases); PFS benefit not specified
	Cost-effectiveness	Chronic phase: 332 000 SEK/QALY; QALY benefit: 3.45 Accelerated phase: 216 000–221 000 SEK/QALY; QALY benefit: 0.85/1.03 Blast phase: Dominated; QALY loss: 0.21/0.27	Chronic phase: Dominating to £23 000/QALY; QALY benefit: 2.54–5.32 Accelerated phase: Dominating to £16 000/QALY; QALY benefit: 0.78–2.6	Chronic phase: £18 000–£37 000/QALY; QALY benefit not specified Accelerated phase: Dominating to £62 000/QALY; QALY benefit not specified

(*continues*)

Table 3

(Continued)

		TLV	SMC	NICE
		ALL: Dominating to 329 000 SEK/QALY; QALY benefit: 0.35/0.36	Blast phase: Dominating to dominated; QALY change: −0.26–0.29 ALL: Dominating to £24 000/QALY; QALY benefit: 0.35	Blast phase: Dominating to £21 000/QALY; QALY benefit not specified ALL: Dominating to £30 000/QALY; QALY benefit not specified
	Decision aspects	Decision modifiers: Moderate to high severity of disease, small population size, limited BIM.	Decision modifiers: Substantial improvement in life expectancy, small population size.	Decision modifier: End of life in accelerated and blast phases of CML as well as ALL
Idelalisib	Decision	Reimbursed	Accepted for use within PAS	Not applied
	Treatment effect	*Vs*. nonresponders in pivotal study; PFS benefit: 4.5 months	*Vs*. prior treatment; PFS benefit not specified	
	Cost-effectiveness	<1 200 000 SEK/QALY; QALY benefit not specified	£62 000/QALY without discount; QALY benefit: 0.35	
	Decision aspects	Decision modifier: Small population size.	Decision modifiers: End of life, small population size.	
Ceritinib	Decision	Reimbursed with risk-sharing agreement	Accepted for use	Recommended within PAS
	Treatment effect	Indirect comparison (partly matched) and *vs*. nonresponders in pivotal study; PFS benefit not explicit (expressed as 5.7–7.2 *vs*. 1.8–3.0 months)	Indirect comparison and *vs*. nonresponders in pivotal study; PFS benefit not specified	Indirect comparison; PFS benefit: 5.2 months
	Cost-effectiveness	440 000–930 000 SEK/QALY; QALY benefit not specified	£50 000/QALY; QALY benefit not specified	<£50 000/QALY with discount; QALY benefit: 0.80
	Decision aspects	Decision modifier: very high severity of disease. Conditioned: Updated health economic analysis required when results from a phase III trial are available.	Decision modifiers: end of life, substantial improvement in life expectancy, small population size.	Decision modifiers: end of life, substantial improvement in life expectancy, small population size, limited BIM.

AL, anaplastic lymphoma; ALL, acute lymphatic leukaemia; BIM, budget impact; CML, chronic myeloid leukaemia; CLL, chronic lymphocytic leukaemia; FL, follicular lymphoma; HL, Hodgkin lymphoma; ICER, incremental cost effectiveness ratio; N/A, not applicable; NICE, National Institute for Health and Care Excellence; PAS, Patient Access Scheme, PFS, progression-free survival; QALY, quality-adjusted life year; SCT, stem cell transplantation; SEK, Swedish kronor; SMC, Scottish Medicines Consortium; TLV, Swedish Dental and Pharmaceutical Benefits Agency.

survival, another surrogate endpoint frequently used in oncology trials, the correlation with overall survival has also been shown to be limited, especially concerning the hazard ratio [14, 15]. Furthermore, as shown in recent reviews regarding drugs approved on the basis of limited evidence, information on clinical outcomes are seldom provided in the postapproval period [16–18]. However, it also needs to be acknowledged that it is difficult to estimate overall survival reliably for cancer diseases where the patients can be expected to live considerably longer than the study duration.

The fact that prospective case series cannot be blinded poses an important limitation related to the design, in particular when the assessments of the primary outcome involve subjectivity. Regarding the subjectivity of the primary outcomes in this review, the inter-rater agreement was acceptable in most studies. To illustrate the extent of subjectivity, a κ value of 0.68 and 100 patients evaluated implies that the assessors assess 16 patients differently. With a κ of 0.24, the assessors' assessments differ in 38 out of 100 patients. Therefore, the subjectivity of the assessment of the primary outcome cannot be neglected and needs to be considered in the interpretation of the results. Indeed, in single-armed trials both investigators and independent review committees are aware that all patients receive the new treatment. Further, our finding that information regarding the patient recruitment process was not provided in the pivotal studies may add to the difficulties in determining the validity of the reported treatment effects. In fact, without this information convenience sampling cannot be excluded.

Regarding the challenge of identifying an appropriate comparison to estimate the size of the drug treatment effect, all three approaches identified in this review have limitations. Indirect comparisons, the most frequently used method, imply that results from studies investigating the effects of other drugs are used to inform decision making. This approach has the obvious limitation that patients from different populations with different characteristics are compared with one another, including factors of importance for the disease prognosis and studied outcomes. This is particularly prominent in naïve comparisons but may also be a substantial problem in matched comparisons. Thus, the estimations of drug treatment effect sizes and subsequent cost-

effectiveness estimates based on single-arm studies are almost inherently associated with huge uncertainties. This may be exemplified by the ponatinib case (Table 3), where gains of up to 5.32 QALYs were estimated. As only one of 18 deaths in the pivotal study was assessed to be caused by disease progression, the resulting gain in health may be surprising.

When using outcome of prior treatment as comparator to determine the effect size, it should be noted that the way to obtain data may differ between the compared groups. Indeed, in one reviewed case, data on effects of the new drug were collected systematically and prospectively, while data on prior treatment were collected in a more *ad hoc* manner and were restricted to information available in the medical records.

Using nonresponders in the pivotal study as comparator may entail that patients with less favourable prognosis constitute the comparison group. This approach may therefore result in an overestimation of effect size and, ultimately, an underestimation of cost per QALY gained. On the other hand, nonresponders may also have responded to some extent, although this may not have been captured in a dichotomized outcome. The method may therefore also hold the opposite risk, with an overestimated cost per QALY gained.

For new drugs addressing unmet medical needs, prevailing expectations that register data will provide information on drug effectiveness needed for decision making may have contributed to the introduction of accelerated approval processes [3]. However, the results of a recent review of published effectiveness studies based on Swedish Prescribed Drug Register data [19] may hamper these expectations as few studies passed an overall quality assessment of the evidence achieved, and only one out of 24 publications had a low risk of bias [20]. Further, a review of drug approvals by the Food and Drug Administration illustrates that randomized controlled trials can be performed also for rare cancer diagnoses [21].

When quantifying potential gains in health, adverse reactions need to be considered. Indeed, as all incurable cancers can be regarded as presenting unmet needs when it comes to treatment, the benefit–risk balance is crucial when determining whether the new drug can meet an unmet medical need and consequently be approved. In this review, all drugs were intended for severe cancer. Importantly, and relevant for the benefit–risk assessment, all drugs were also shown to have severe side effects. The fact that, according to investigator assessments, 11 deaths in the pivotal studies were due to disease progression and six were related to treatment may deserve some attention. Indeed, two in six deaths for brentuximab vedotin, three in six for bosutinib, seven out of seven for vismodegib, 17 in 18 for ponatinib and eight in 11 for idelalisib, occurring on treatment or within 30 days of study drug discontinuation, were not due to disease progression, and adverse reactions cannot be excluded. Assessments of causality are difficult, especially when the adverse reactions are unknown, as for new drugs, or when the reaction is a common condition, for example myocardial infarction or stroke. Therefore, the investigators' assessments of causality need to be interpreted with caution and the reported number of deaths related to treatment may be underestimated. Although our findings may suggest severe risks with the reviewed new drugs, the EMA assessed the risks as acceptable or manageable given the benefits which they considered established by the one-armed trials. Interestingly, postmarketing safety problems have been found more common for drugs granted accelerated approval than drugs given regular approval [21, 22]. Further, it may be worth mentioning that data on HRQL were only recorded in two of the pivotal studies, and in none with the EQ-5D instrument, which is often preferred by the reimbursement agencies. Such data may provide valuable insights for the benefit–risk balance assessments as well as for the cost-effectiveness analyses. In this review, the latter frequently relied on HRQL data from nonpivotal trial sources.

For value-based pricing and resource allocation in publicly funded health care, transparency is crucial. Interestingly, this review illustrates that drug treatment effect sizes and estimated gains in health were not systematically specified in the decisions/guidance documents. To facilitate the understanding for health care decision makers and providers, including health care personnel, it may be of value to clearly state these figures. As the extensive information by the Evidence Review Group providing assessments for decision making in the NICE was not reviewed, we cannot exclude that this information is available in their documentation. Regarding cancer treatments, the efforts by the European Society for Medical Oncology to explicitly assess the beneficial effect of new cancer treatments constitute an example of an important step in the transparency direction [23]. Indeed, less than one third of contemporary randomized controlled trials with statistically significant results met the European Society for Medical Oncology thresholds for meaningful clinical benefit [24].

Commendably, the uncertainties associated with the reviewed decisions and advice/guidance documents were discussed in all cases. Although acknowledged and elaborated on, the fact that most of the new drug treatments were allowed for use in public health care indicates that the reimbursement agencies considered the underlying documentation sufficient for decision making. However, it should be noted that unofficial discounts, attached to several granting decisions, may be an important factor in the handling of uncertainty in the reimbursement decisions. Further, some decisions were conditioned by follow-up requirements. In addition, the decision modifiers mentioned earlier imply that a larger degree of uncertainty in the evidence base may be accepted in the reviewed cases compared with cases where these modifiers do not apply. A recently suggested approach may be to use price as a lever for reimbursement explicitly linking acceptable price to the extent of uncertainty in evidence [25]. However, as many new drugs, at least in this review, have a real potential to harm, a direct link to pricing may not be readily made.

To illustrate the market value of the reviewed drugs, sales statistics from the Swedish eHealth Agency, covering all dispensed prescription drugs from every pharmacy in Sweden (approximately 10 million inhabitants), reveal that the health care costs for the reviewed drugs (including idelalisib used in chronic lymphatic leukaemia, which cannot be separated from follicular lymphoma in the sales statistics) amounted to 32.8 million Swedish kronor in 2016, corresponding to 3.5 million Euro. Given a similar magnitude of use in the UK including Scotland (approximately 65 million inhabitants), these costs may amount to 26 million Euro. This figure can be expected to grow in the future,

emphasizing the importance of thoughtful decision making for drugs with a sparse evidence base.

Strengths and limitations

An important strength of the present study is that it provides information on how six challenging and current cases were handled by established European reimbursement agencies. Starting from a selection of new drugs, requested for use in public health care by professional expert groups, our review should reflect clinically relevant cases. This review may therefore facilitate discussions on future handling of challenges associated with drug market approvals based on prospective case series. The low number of available cases needs to be recognized and constitutes a limitation in terms of generalizability. However, the number of cases with evidence restricted to prospective case series is likely to grow, and a review such as the present one may contribute to lessons for the future as every reimbursement decision may be considered a precedent case.

Conclusions and implications

In summary, we show that drugs with market approvals based on prospective case series are granted reimbursement in most cases, although the underlying evidence has limitations regarding the internal and the external validity over and above the design *per se*. This calls for increased methodological rigour in the design and reporting of such studies. Further, our review elucidates that the cost-effectiveness analyses are hugely uncertain because of difficulties to quantify drug treatment effect sizes. Therefore, as early access of new drugs is a pursuit of today and transparency desirable, further efforts in clarifying the role of different decision modifiers are essential for value-based pricing decision making.

Competing Interests

All authors have completed the Unified Competing Interest form (available on request from the corresponding author) and declare: a grant from the Swedish Dental and Pharmaceutical Benefits Agency (TLV) to provide a report on the topic of this manuscript; being former (S.M.W.) and present (M.H.) members of the Pharmaceutical Benefits Board at TLV deciding on reimbursement status regarding prescription drugs; no other relationships or activities that could appear to have influenced the submitted work.

The study was funded by ALF grants (ALFGBG-428711, ALFGBG-716941) and the Swedish Dental and Pharmaceutical Benefits Agency (TLV). The funding sources did not influence the design, methods, analysis, or preparation of the paper, or the decision to submit the paper for publication.

Contributors

S.M.W. conceived the study. S.M.W. and M.H. designed the study and extracted the data. S.M.W. performed the statistical analyses and drafted the manuscript. M.H. revised the manuscript for intellectual content. S.M.W. and M.H. are the guarantors of this work.

References

1 Wallerstedt SM. Medical and scientific assessments as the basis for prioritization and resource allocation for new drug treatment – experiences from five years of systematic work [Swedish]. Lakartidningen 2016; 113: D43D.

2 Hatswell AJ, Baio G, Berlin JA, Irs A, Freemantle N. Regulatory approval of pharmaceuticals without a randomised controlled study: analysis of EMA and FDA approvals 1999-2014. BMJ Open 2016; 6: e011666.

3 Martinalbo J, Bowen D, Camarero J, Chapelin M, Demolis P, Foggi P, *et al*. Early market access of cancer drugs in the EU. Ann Oncol 2016; 27: 96–105.

4 European Medicines Agency. Conditional marketing authorisation – report on ten years of experience at the European Medicines Agency. 2017. Available at http://www.ema.europa.eu/docs/en_GB/document_library/Report/2017/01/WC500219991.pdf (last accessed 10 January 2018).

5 Regional Health Technology Assessment Centre. Checklist for case series [Granskningsmall för fallerier]. 2017. Available at https://www2.sahlgrenska.se/sv/SU/Forskning/HTA-centrum/Hogerkolumn-undersidor/Hjalpmedel-under-projektet/ (last accessed 10 January 2018).

6 Harding SD, Sharman JL, Faccenda E, Southan C, Pawson AJ, Ireland S, *et al*. The IUPHAR/BPS guide to PHARMACOLOGY in 2018: updates and expansion to encompass the new guide to IMMUNOPHARMACOLOGY. Nucl Acids Res 2018; 46: D1091–106.

7 Sekulic A, Migden MR, Oro AE, Dirix L, Lewis KD, Hainsworth JD, *et al*. Efficacy and safety of vismodegib in advanced basal-cell carcinoma. N Engl J Med 2012; 366: 2171–9.

8 Younes A, Gopal AK, Smith SE, Ansell SM, Rosenblatt JD, Savage KJ, *et al*. Results of a pivotal phase II study of brentuximab vedotin for patients with relapsed or refractory Hodgkin's lymphoma. J Clin Oncol 2012; 30: 2183–9.

9 Gopal AK, Kahl BS, de Vos S, Wagner-Johnston ND, Schuster SJ, Jurczak WJ, *et al*. PI3Kdelta inhibition by idelalisib in patients with relapsed indolent lymphoma. N Engl J Med 2014; 370: 1008–18.

10 Cortes JE, Kim DW, Pinilla-Ibarz J, le Coutre P, Paquette R, Chuah C, *et al*. A phase 2 trial of ponatinib in Philadelphia chromosome-positive leukemias. N Engl J Med 2013; 369: 1783–96.

11 Cortes JE, Kantarjian HM, Brummendorf TH, Kim DW, Turkina AG, Shen ZX, *et al*. Safety and efficacy of bosutinib (SKI-606) in chronic phase Philadelphia chromosome-positive chronic myeloid leukemia patients with resistance or intolerance to imatinib. Blood 2011; 118: 4567–76.

12 Prasad V, Kim C, Burotto M, Vandross A. The strength of association between surrogate end points and survival in oncology: a systematic review of trial-level meta-analyses. JAMA Intern Med 2015; 175: 1389–98.

13 Petrelli F, Barni S. Surrogate endpoints in metastatic breast cancer treated with targeted therapies: an analysis of the first-line phase III trials. Med Oncol 2014; 31: 776.

14 Liu L, Chen F, Zhao J, Yu H. Correlation between overall survival and other endpoints in metastatic breast cancer with second- or third-line chemotherapy: literature-based analysis of 24 randomized trials. Bull Cancer 2016; 103: 336–44.

15 Shitara K, Matsuo K, Muro K, Doi T, Ohtsu A. Correlation between overall survival and other endpoints in clinical trials of second-line chemotherapy for patients with advanced gastric cancer. Gastric Cancer 2014; 17: 362–70.

16 Pease AM, Krumholz HM, Downing NS, Aminawung JA, Shah ND, Ross JS. Postapproval studies of drugs initially approved by the FDA on the basis of limited evidence: systematic review. BMJ 2017; 357: j1680.

17 Davis C, Naci H, Gurpinar E, Poplavska E, Pinto A, Aggarwal A. Availability of evidence of benefits on overall survival and quality of life of cancer drugs approved by European Medicines Agency: retrospective cohort study of drug approvals 2009-13. BMJ 2017; 359: j4530.

18 Naci H, Smalley KR, Kesselheim AS. Characteristics of preapproval and postapproval studies for drugs granted accelerated approval by the US Food and Drug Administration. JAMA 2017; 318: 626–36.

19 Wallerstedt SM, Wettermark B, Hoffmann M. The first decade with the Swedish prescribed drug register – a systematic review of the output in the scientific literature. Basic Clin Pharmacol Toxicol 2016; 119: 464–9.

20 Wallerstedt SM, Hoffmann M. Evaluating beneficial drug effects in a non-interventional setting: a review of effectiveness studies based on Swedish prescribed drug register data. Br J Clin Pharmacol 2017; 83: 1309–18.

21 Gaddipati H, Liu K, Pariser A, Pazdur R. Rare cancer trial design: lessons from FDA approvals. Clin Cancer Res 2012; 18: 5172–8.

22 Downing NS, Shah ND, Aminawung JA, Pease AM, Zeitoun JD, Krumholz HM, *et al*. Postmarket safety events among novel therapeutics approved by the US Food and Drug Administration between 2001 and 2010. JAMA 2017; 317: 1854–63.

23 Cherny NI, Sullivan R, Dafni U, Kerst JM, Sobrero A, Zielinski C, *et al*. A standardised, generic, validated approach to stratify the magnitude of clinical benefit that can be anticipated from anti-cancer therapies: the European Society for Medical Oncology Magnitude of Clinical Benefit Scale (ESMO-MCBS). Ann Oncol 2015; 26: 1547–73.

24 Del Paggio JC, Azariah B, Sullivan R, Hopman WM, James FV, Roshni S, *et al*. Do contemporary randomized controlled trials meet ESMO thresholds for meaningful clinical benefit? Ann Oncol 2017; 28: 157–62.

25 Pace J, Ghinea N, Kerridge I, Lipworth W. Demands for access to new therapies: are there alternatives to accelerated access? BMJ 2017; 359: j4494.

26 Pro B, Advani R, Brice P, Bartlett NL, Rosenblatt JD, Illidge T, *et al*. Brentuximab vedotin (SGN-35) in patients with relapsed or refractory systemic anaplastic large-cell lymphoma: results of a phase II study. J Clin Oncol 2012; 30: 2190–6.

27 Khoury HJ, Cortes JE, Kantarjian HM, Gambacorti-Passerini C, Baccarani M, Kim DW, *et al*. Bosutinib is active in chronic phase chronic myeloid leukemia after imatinib and dasatinib and/or nilotinib therapy failure. Blood 2012; 119: 3403–12.

28 Sekulic A, Migden MR, Basset-Seguin N, Garbe C, Gesierich A, Lao CD, *et al*. Long-term safety and efficacy of vismodegib in patients with advanced basal cell carcinoma: final update of the pivotal ERIVANCE BCC study. BMC Cancer 2017; 17: 332.

29 Shaw AT, Kim DW, Mehra R, Tan DS, Felip E, Chow LQ, *et al*. Ceritinib in ALK-rearranged non-small-cell lung cancer. N Engl J Med 2014; 370: 1189–97.

30 Kim DW, Mehra R, Tan DS, Felip E, Chow LQ, Camidge DR, *et al*. Activity and safety of ceritinib in patients with ALK-rearranged non-small-cell lung cancer (ASCEND-1): updated results from the multicentre, open-label, phase 1 trial. Lancet Oncol 2016; 17: 452–63.