Annals of Oncology Advance Access published July 26, 2016

original article

Annals of Oncology 00: 1–8, 2016
doi:10.1093/annonc/mdw233

Relative effectiveness assessments of oncology medicines for pricing and reimbursement decisions in European countries

S. Kleijnen[1,2*], I. Lipska[2,3], T. Leonardo Alves[2], K. Meijboom[4], A. Elsada[5], V. Vervölgyi[6], A. d'Andon[7], A. Timoney[8], H. G. Leufkens[2], A. De Boer[2] & W. G. Goettsch[1,2]

[1]National Health Care Institute, Diemen, The Netherlands; [2]Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences, Utrecht, The Netherlands; [3]Center for Innovation in Regulatory Science, London, UK; [4]VU University Amsterdam, Amsterdam, The Netherlands; [5]National Institute for Health and Care Excellence, Manchester, UK; [6]Department of Drug Assessment, Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, Cologne, Germany; [7]Medicines Evaluation Department, Haute Autorité de santé, Paris, France; [8]NHS Lothian and Strathclyde Institute of Pharmacy and Biomedical Sciences, University of Strathclyde, Glasgow, UK

Received 17 December 2015; revised 6 April 2016 and 1 June 2016; accepted 1 June 2016

Background: There is a debate on the added clinical value of new, expensive, anticancer treatments. Among European decision makers, the relevance of commonly used end points in trials, especially overall survival (OS), progression-free survival (PFS) and quality of life (QoL), varies, leading to the available evidence being valued differently. This research studies the extent to which the value of end points for cancer medicines differs among European decision makers.

Methods: We compared guidelines and relative effectiveness assessments (REAs) of medicines for pricing or reimbursement decisions in England, France, Germany, The Netherlands, Poland, and Scotland. Anticancer medicines that received marketing authorization in Europe between 2011 and 2013 with at least four available national REAs were evaluated. A total of 79 REAs were included.

Results: Health technology assessment (HTA) guidelines indicate a preference for clinically and patient relevant end points such as OS and QoL above surrogate end points. Most guidelines do not specify whether PFS is considered a surrogate or patient-relevant end point. The number of REAs included per jurisdiction varied between 7 (The Netherlands) and 18 (Germany). OS data were included in all REAs and were the preferred end point by HTA agencies, but these data were not always mature or robust. QoL data are included in only 54% of the REAs, with a limited impact on the recommendations. PFS data are included in 70% of the REAs, but the extent to which HTA agencies find PFS relevant varies.

Conclusion(s): European decision-making on relative effectiveness of anticancer medicines is affected by a gap in requested versus available clinical evidence, mainly because the regulator is willing to accept some degree of clinical uncertainty. A multi-stakeholder debate would be essential to align concrete robust evidence requirements in oncology and a collectively shared definition for relevant clinical benefit, which will benefit patients and society in general.

Key words: comparative effectiveness, health technology assessment, reimbursement, antineoplastic agents, clinical oncology

introduction

New anticancer medicines promise an improved prognosis for patients with life-threatening diseases. However, most of them are modestly effective while very expensive [1]. This dilemma frequently leads to a multi-stakeholder debate about the value of such medicines entering the European market.

Unlike the centralized European marketing authorization decision, each member state independently makes its own reimbursement decisions. One of the most important criteria for reimbursement decisions is usually the comparative efficacy and/or effectiveness of the new treatment with existing options [2, 3]. This comparison is often referred to as a relative efficacy/effectiveness assessment (REA) and is carried out by health technology assessment (HTA) agencies. Other relevant factors in reimbursement decisions include ethical, social, budget-impact, and cost-effectiveness considerations [3].

Evidence from case studies suggests that how HTA agencies value commonly used clinical end points for anticancer medicines differs [4, 5]. Generally, the potential benefits of a new treatment come down to its effect on overall survival (OS) and/or quality of life (QoL), or their surrogates [6]. Examples of

Correspondence to: Mrs Sarah Kleijnen, Heath Care Department, Zorginstituut Nederland, Eekholt 4, 1112 XH Diemen, The Netherlands. Tel: +31-207978008; E-mail: skleijnen@zinl.nl

surrogates are disease-free survival in the curative setting, and progression-free survival (PFS) in the non-curative setting.

PFS is usually defined as the time from random assignment in a clinical trial to tumor progression or death from any cause. The increasing use of PFS as a primary end point in anticancer trials is debated due to doubts about its clinical meaningfulness [7, 8]. Advanced colorectal and advanced ovarian cancer seem to be the only two tumor types for which evidence suggests that PFS may be accepted as surrogate for OS [7]. But even for these indications, the validity of this association in contemporary oncology with novel therapies is being questioned [7]. It is also debated whether PFS can measure a direct clinical benefit in the advanced setting [9]. This would be the case if it provides a duration in which patients experience less symptoms, clinical consequences of the disease, and/or improved QoL [5].

A comparison of pazopanib assessments for advanced/metastatic renal cell carcinoma, found that some HTA agencies considered an increase in PFS to be patient relevant, whereas other agencies considered it only relevant in the absence of OS data and when supported by improved QoL [4]. In addition, a study comparing appraisals of breast cancer and colorectal cancer medicines across five HTA agencies found that HTA agencies interpreted the PFS benefit differently [5].

Greater harmonization in assessing clinical end points for anticancer medicines is important to patients, healthcare providers, and payers to guide appropriate treatment decisions. The objective of this research is to study the role of OS, PFS, and QoL data in REAs informing pricing or reimbursement decisions in European jurisdictions, by (i) studying whether data on these end points are included, and (ii) studying the impact of these data on recommendations.

methods: research design

We conducted a retrospective comparative cross-sectional analysis of publicly available HTAs of anticancer medicines that received marketing authorization between 2011 and 2013.

selection of HTA jurisdictions

Of the 29 EU jurisdictions (UK divided into England and Scotland), nine had publicly available reports from HTA organizations involved in assessing medicines for pricing or reimbursement decisions. From these, three were excluded. Belgium was excluded as only a limited number of reports were publicly available; Portugal and Ireland were excluded as only a few brief summaries were available, providing insufficient information to inform this study. We present the six jurisdictions included and their respective HTA agencies in Table 1.

HTA guidelines

National HTA guidelines assessing medicines were obtained from the relevant HTA agencies' websites. If no national guideline was available, grey literature was searched.

Table 1. Overview of health technology assessment agencies included in the study

Jurisdiction	England	France	Germany	The Netherlands	Poland	Scotland
HTA organization	NICE	HAS	IQWiG	ZIN	AOTMiT	SMC
Section in report about relative effectiveness assessment	Evidence for clinical effectiveness[a]	Clinical added value (ASMR)	Early benefit assessment	Pharmaco therapeutic assessment	Clinical efficacy[a]	Clinical effectiveness issues[a]
Section in report used to identify impact of end point on recommendation	Summary of Appraisal Committee's key conclusions	Transparency Committee conclusions and summary and discussion section of the analysis of data	Executive summary of added benefit assessment and extent and probability of added benefit	Letter to Ministry of Health, Welfare and Sports and Judgement medicines committee discussion/ extrapolation/ conclusion section	Justification of recommendation	Advice and summary of clinical effectiveness issues
Other criteria (besides relative effectiveness) taken into account by agency in recommendation	Cost-effectiveness	Cost-effectiveness for claimed ASMR I–III and/or high budget-impact	None	Cost-effectiveness for premium pricing, budget impact	Cost-effectiveness, budget-impact	Cost-effectiveness
Recommendation is used as basis for	Funding decision	Reimbursement and pricing decision (positive list)	Pricing decision (reference pricing, price negotiations with social insurance)	Reimbursement and pricing decision[b]	Reimbursement and pricing decision (positive list)	Funding decision by Health Boards

AOTMiT, Agencja Oceny Technologii Medycznych i Taryfikacji; HAS, Haute Autorité de Santé; HTA, Health technology assessment; IQWiG, Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen; MA, Market authorization; NICE, National Institute for Health and Care Excellence; REA, relative effectiveness assessment; SMC, Scottish Medicines Consortium; ZIN, Zorginstituut Nederland.

[a]No separate REA recommendation.

[b]Pricing decision is only applicable to outpatient medicines, for which a positive list is in place.

selection of medicines and reports

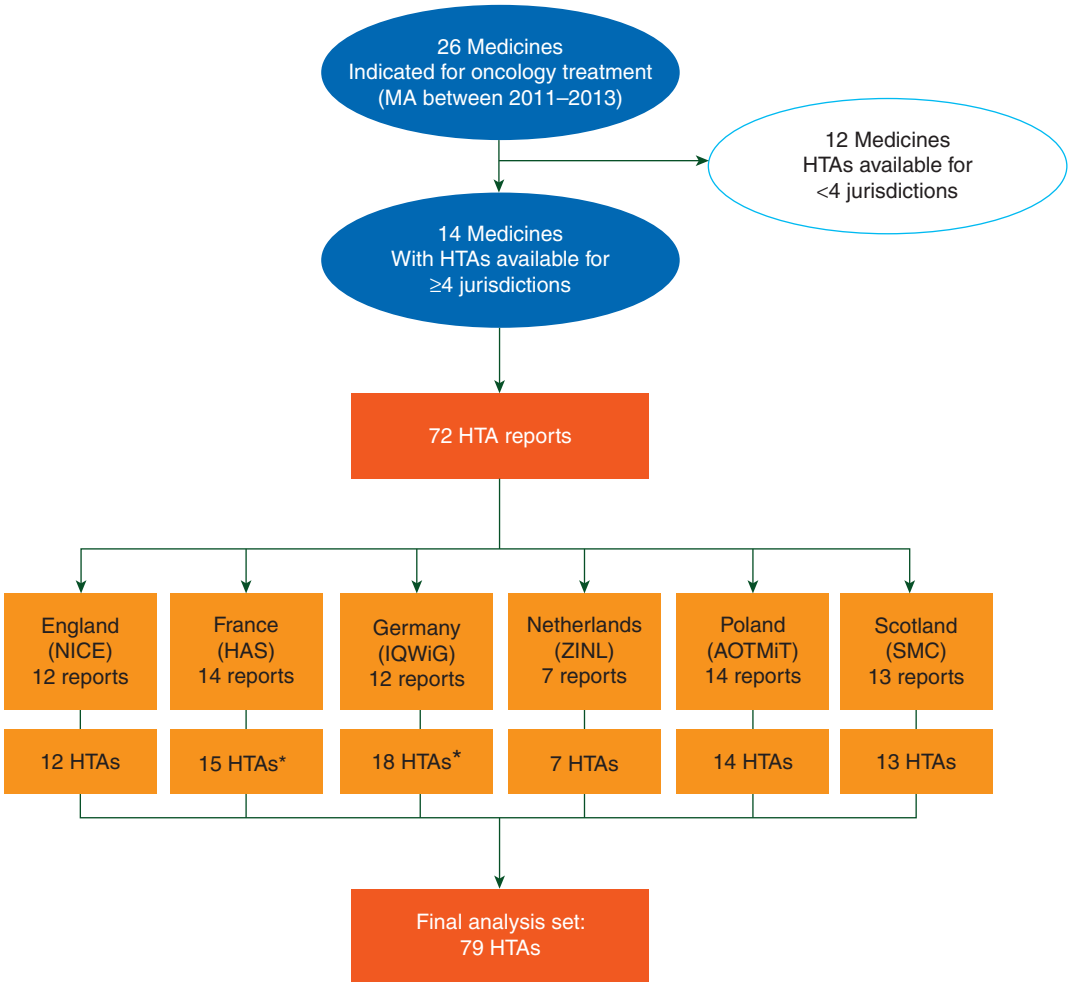
Of all new active substances approved by the EMA from 1 January 2011 to 31 December 2013 to treat malignant diseases ($n = 26$), we included only those medicines for which four or more HTA reports were published before April 2015 by different HTA agencies for the first indication approved ($n = 14$). A total of 72 HTA reports for these 14 medicines were included. When an HTA report included separate analyses and/or recommendations for individual (sub)indications, we included each (sub)indication separately. Although the 12 reports from Germany's Institute for Quality and Efficiency in Health Care (IQWiG) included 25 (sub)indications with separate recommendations, we excluded 7 indications because data were missing, resulting in $n = 18$. One report from France's National Authority for Health (HAS) included two indications with separate recommendations. The final dataset included 79 HTAs. We present a flowchart of the selection process in Figure 1.

data collection and analysis

To collect data from the assessments, we developed a structured data collection form (DCF) including 32 questions, 14 open-ended, and 18 categorical. The DCF and a description of its development are presented in supplementary Table S1 and S2, available at *Annals of Oncology* online. This article focuses on a subset of questions in the DCF that are related to the research questions.

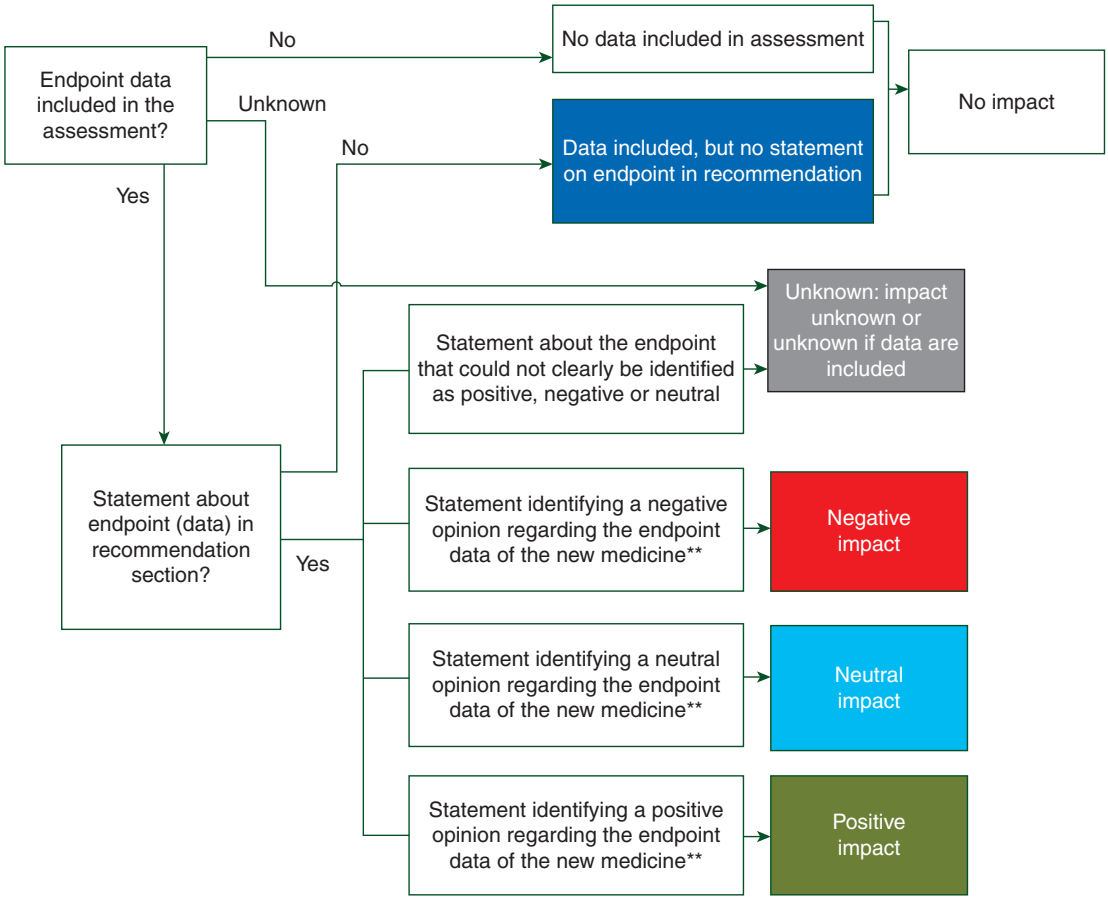
As the study focuses on relative effectiveness rather than cost-effectiveness, we extracted statements about the end points from the clinical sections of the reports and from the overall recommendations or discussion sections (Table 1). QoL data had to be collected with validated QoL instruments. To capture the impact of the clinical end points on the recommendations, we categorized the extracted statements as *positive, neutral, negative, unknown* (impact unknown or unknown if data are included), or *no impact* (not included/not identified). Statements were classified as neutral if it indicated that no change/difference is shown versus comparator. We present the algorithm for the categorization in Figure 2.

We abstracted data between April and May 2015 and invited an expert panel consisting of one representative from each of the six agencies who are or have been involved in producing HTAs. Their role was to validate the algorithm used to categorize the impact of the end point and to clarify pending issues.



*When an HTA report included separate analysis and/or recommendations for several (sub)indications, each (sub)indication was included as a separate assessment.
Abbreviations: AOTMiT = Agencia Oceny Technologii Medycznych i Taryfikacji; HAS = Haute Autorité De Santé: HTA = Health Technology assessment; IQWiG = Instltut für Qualität und Wirtschaftlichkeit Im Gesundheitswesen; MA = Market Authorisation; NICE = National Institute for Health and Care Excellence: REA = Relative Effectiveness Essessment: SMC = Scottish Medicines Consortium: ZIN = Zorginstituut Nederland.

Figure 1. Flow chart: selection of medicines and health technology assessments.



*The impact was classified as unknown in case of multiple comparators with different impact values and it was not possible to choose a single most relevant comparator (e.g. England, axitinib and afatinib). In addition, for some Polish reports it is unknown whether endpoint data are included due to confidential (sensored) sections.

**Based on direct statement in recommendation/discussion on endpoint OR indirect statement (e.g superior efficacy) that is clearly related to a specific endpoint

Figure 2. Algorithm used to determine the impact of the end point data on recommendation.

We used descriptive statistics to present the data and qualitatively analyzed statements to compare what agencies regard as clinically relevant in the context of trial end points, and why these judgments may differ.

results

HTA guidelines

Information in the guidelines on end points is presented in supplementary Table S3, available at *Annals of Oncology* online. In general, all HTA guidelines preferred clinically and patient-relevant end points relating to morbidity, mortality, and QoL. Surrogate end points are not favored, but used when supporting information is provided about the relationship between the surrogate and patient-relevant end points. Most guidelines do not specify whether PFS is considered a surrogate or patient-relevant end point. A French consensus statement by clinical experts indicates that PFS in metastatic disease is relevant only in certain settings. On the contrary, a German report on surrogate end points in oncology concluded that PFS should not be considered a valid surrogate for OS in colorectal and breast cancer.

The guidelines from England and Scotland make special provisions for life-extending treatments at the end of life, which can result in a higher valuation of the clinical benefit offered by treatment reflected in a higher cost-effectiveness threshold.

HTAs included and recommendation outcomes

We list all assessed medicines and recommendation outcomes in Table 2. The number of HTAs included per jurisdiction varied between 7 (The Netherlands) and 18 (Germany). Twenty-seven percent (21/79) of the assessments had a negative/lesser benefit recommendation, but the percentage varies considerably per jurisdiction (6%–69%). Overall, few medicines were rejected primarily for clinical reasons (4/79 recommendations), whereas 10 of 79 were rejected primarily because of cost/cost-effectiveness issues. For 7 of 79 assessments, the rejection was based on the clinical and cost/cost-effectiveness profile. For France, Germany, and The Netherlands, negative/lesser benefit recommendations were based solely on the clinical profile, whereas they were mainly based on the cost/cost-effectiveness profile or both (clinical and cost/cost-effectiveness profile) for England, Scotland, and Poland.

Table 2. List of medicines included and outcome of recommendations that inform pricing and/or reimbursement decisions

Abbreviated indication	Medicine (generic name)	England	France	Germany	The Netherlands	Poland	Scotland
Bone metastases from solid tumors	Denosumab	⊕ (optimized)	⊕ (minor)[a] ⊕/⊖[a]	Not assessed	⊕/⊖	⊖, c and €	Not assessed
Breast cancer	Eribulin	⊖, €	⊕ (minor)	⊕/⊖[b] ⊕/⊖[b]	⊕	⊖, c and €	⊖, €[c]
	Pertuzumab	Not assessed	⊕ (moderate)	⊕ (major)[d]	Not assessed	⊕	⊖[e], €
Colorectal cancer	Aflibercept	⊖, c and €	⊕/⊖	⊕ (minor)	Not assessed	⊕	⊖[f], €
Gastric cancer	Tegafur/gimeracil/ oteracil	Not assessed	⊖, c	Not assessed	⊖, c	⊖, c	⊕ (with restrictions)
Melanoma	Ipilimumab, second-line Tx	⊕[g]	⊕ (minor)[h]	⊕ (considerable)	⊕	⊕	⊖, c and €
	Vemurafenib	⊕[g]	⊕ (moderate)	⊕ (considerable)[i]	⊕	⊕	⊖, €[c]
	Dabrafenib	⊕	⊕/⊖	⊕	Not assessed	⊕	⊕ (with restrictions)[j]
Non-small-cell lung cancer	Afatinib	⊕	⊕/⊖	⊕ (major)[k] ⊕ (minor)[k] ⊕/⊖[k] ⊖, c[k]	Not assessed	⊕	⊕
	Crizotinib	⊖, €[g]	⊕ (moderate)	⊕/⊖	Not assessed	⊖, c and €	⊖, €[c]
Prostate cancer	Cabazitaxel	⊖, €	⊕ (minor)[l]	⊕ (considerable)[m] ⊕ (minor)[m]	⊕	⊖, c and €	⊖, c and €
	Abiraterone, after Tx with taxane	⊕[g]	⊕ (moderate)	⊕ considerable)[n]	⊕/⊖	⊕	⊖ €[o]
	Enzalutamide	⊕[g]	⊕ (moderate)	⊕ (considerable)[p] ⊕ (major)[p]	Not assessed	⊕	⊕ (with restrictions)
Renal-cell carcinoma	Axitinib	⊕ (optimized)[g]	⊕ (minor)	⊕ (considerable)[q]	Not assessed	⊕	⊖ €[c]
	# assessments	12	15	18	7	14	13
	n ⊖/% ⊖	4/33%	1/7%	1/6%	1/14%	5/36%	9/69%

⊕, recommended/added benefit; ⊕/⊖, no added benefit proven (GE)/similar therapeutic value (NL, FR); ⊖, not recommended (EN, PO, SC)/lesser benefit (FR, GE, NL); c, clinical profile (benefit, harms) was the primary reason for negative recommendation; €, costs/cost-effectiveness was the primary reason for negative recommendation; c and €, clinical profile (benefit, harms) and costs/cost-effectiveness were the primary reason for negative recommendation; Tx, treatment.

[a]HAS recommended that denosumab provides a minor improvement in actual benefit (level IV) in patients with breast cancer or prostate cancer with bone metastases but does not provide an improvement in actual benefit (level V) in patients with other types of solid tumors with bone metastases.

[b]Eribulin was assessed for two subgroups: patients for whom treatment with taxanes or anthracyclines is no longer an option versus patients for whom repeated treatment containing an anthracycline or a taxane is an option.

[c]The medicine was reassessed by SMC for the same indication. The reassessment included the same clinical data; however, the recommendation changed from negative to positive with limitations due to a Patient Access Scheme that improved the cost-effectiveness of the medicine.

[d]IQWIG assessed two separate subpopulations of which only one was included in the analysis: HER2-positive metastatic breast cancer (hint of major added benefit).

[e]The same indication was reassessed in 2014, for which an updated efficacy analysis was added. But the recommendation remained negative due to an insufficiently robust economic analysis and the high treatment cost in relation to the health benefits.

[f]The medicine was reassessed by SMC for the same indication. The reassessment included the same clinical data. However, the recommendation changed from negative to positive due to a Patient Access Scheme that improved the cost-effectiveness of the medicine.

[g]NICE end-of-life criteria were met (life expectancy <24 months, extension to life of at least an additional 3 months, and ≤7000 patients in England). The criteria were introduced in 2009.

[h]Ipilibumab was assessed by HAS in 2011 and 2012 for the same indication, both assessments resulting in a minor incremental added benefit recommendation.

[i]Vemurafenib was assessed by IQWIG in 2012 and 2013 for the same indication both resulting in a considerable added benefit recommendation. For the reassessment, the manufacturer submitted additional data cut-offs; however, the data were not included by IQWIG due to increased risk of bias.

[j]Meets SMC end-of-life criteria (life expectancy ≤36 months). The criteria were introduced in 2014.

[k]Afatibin was assessed by IQWIG for six different subpopulations of which four were included in the analysis: (i) non-pretreated patients with ECOG PS 0–1 and an EGFR mutation Del19 (Indication of a major AB), (ii) non-pretreated patients with ECOG PS 0–1 and an EGFR mutation L858R, age <65 (Hint of a minor AB), (iii) non-pretreated patients with ECOG PS 0–1 and an EGFR mutation L858R, age ≥65 (AB not proven), (vi) non-pretreated patients with ECOG PS 0–1 and other EGFR mutations (indication of LB).

[l]Cabazitaxel was reassessed by HAS for the same indication in 2012. New data were included in the reassessment (2012) resulting in a change of the recommendation from minor improvement in actual benefit to moderate improvement in actual benefit.

[m]Cabazitaxel was assessed by IQWIG for three different subpopulations: (i) best supportive care population, age <65 year (considerable added benefit), (ii) best supportive care population, age ≥65 years (hint of added benefit), (iii) docetaxel retreatment population. The third subpopulation was excluded from the dataset as IQWIG did not conduct an analysis because no data were available.

[n]IQWIG assessed two separate subpopulations of which one was included in the analysis: best supportive care population (considerable added benefit).

[o]Abiterone was assessed twice by SMC for the same indication in 2012, with the same clinical data. However, the recommendation changed from negative to positive due to a Patient Access Scheme that improved the cost-effectiveness of abiterone.

[p]Enzalutamide was assessed by IQWIG for two different subpopulations: (i) patients without visceral metastases (Hint of a major added benefit), (ii) patients with visceral metastases (Hint of a considerable added benefit).

[q]IQWIG assessed two separate subpopulations of which one was included in the analysis: (i) cytokine population (hint of a considerable added benefit).

end point data included in REAs

Figure 3 details the end points included in the REAs and their impact on the recommendations. OS data were included by all agencies in all REAs, but the data are not always mature. Germany did not include PFS data in any of the REAs. In the other jurisdictions, PFS data were included in 80%–100% of the REAs. QoL data are frequently lacking, and inclusion varies, from 29% (Poland) to 67% (England). Where QoL data were not included, this was either because the data were not collected or were immature or not robust. Safety data were included by all jurisdictions for all medicines.

impact of end point data on recommendations

OS and safety data had an impact on the recommendation in 94% and 86% of the HTAs, respectively. The impact of OS data was mainly positive (48%/94%) or neutral (35%/94%), whereas that of safety data was mainly negative (39%/86%) or neutral (34%/86%). PFS data had an impact in 56% of the recommendations, but this varied highly between jurisdictions, from 0% in Germany to 85% in Scotland. The impact of PFS data was mainly positive (35%/56%). The influence of QoL data seems rather limited as only 41% of the recommendations were affected by QoL data, with the impact being mainly neutral (19%/41%) or positive (16%/41%).

In supplementary Table S4, available at *Annals of Oncology* online, we present the impact of the end points for all medicines per jurisdiction in detail. In at least two instances (cabazitaxal and crizotinib), the impact differed between jurisdictions because of how the clinical relevance of the effect size of OS or PFS was interpreted. For example, the effect size of cabazitaxal for prostate cancer (2.4 month OS gain, HR = 0.70) was considered a major added benefit (Germany), or a slight benefit against a high risk of adverse events (Poland). It was explicitly stated in REAs that the PFS gain was considered clinically relevant by multiple jurisdictions for pertuzumab for breast cancer (18.5 versus 12.4 months, HR = 0.62), crizotinib for lung cancer (7.7 versus 3 month, HR = 0.49), vemurafenib for melanoma (5.3versus 1.6 months, HR = 0.26), and afatinib for lung cancer (11.1 versus 6.9 months, HR = 0.58).

discussion

The costs of new anticancer medicines are high, although their clinical value is sometimes disputed [10, 11], resulting in a debate as to whether or not these medicines should be routinely available in public healthcare systems in the EU. Recently, van Harten et al. [12] found that the prices of anticancer medicines varied substantially among 15 European states. Other studies reported that the reimbursement of anticancer medicines varied among European countries [5, 13, 14]. For countries in which healthcare is financed by general taxation, such as the UK, technologies are more likely to be reimbursed the lower their incremental cost-effectiveness ratio is [13]. However, cost-effectiveness does not play a role in countries such as Germany, where decisions are solely based on clinical evidence [13]. In France, cost-effectiveness only influences the price, not the reimbursement. Despite cost-effectiveness being the principal driver of decisions in some European countries, the relative effectiveness of a medicine is the most commonly shared decision-making criterion across all countries [3].

This study adds to the existing knowledge by focusing on differences in the assessment of clinical end points in REAs for anticancer medicines across European HTA agencies. It highlights the existing evidence gap between the ideal situation (preferred type of evidence as requested by HTA agencies) and the reality (actual evidence provided). OS and QoL are considered preferred patient-relevant end points, but conclusive data on these end points are not always available (e.g. immature OS or incomplete/no QoL data). Nevertheless, for QoL, the lack of evidence does not seem to negatively impact the recommendations. The cross-country variation we found in valuing clinical end points was most striking for PFS data.

The variation we found in relevance of PFS data reflects the ongoing debate about the increasing reliance on PFS to demonstrate a clinical benefit for regulatory purposes [7, 9]. Granting early access to novel therapies based on PFS data can benefit patients who need life-extending therapies, but this runs the risk of reimbursing therapies that later prove not as effective or safe as initially thought [15]. We were unable to identify a formal position of HTA agencies about the relevance of PFS from the publicly available data, except for Germany where PFS is explicitly considered to be of limited influence [16]. Interestingly, the German position does not lead to more negative recommendations than the other jurisdictions.

For the other jurisdiction, the HTA guidelines suggest that PFS is generally seen as a surrogate end point, which confirm previous research [5, 17, 18]. But as the HTA agencies are reluctant to discard the data despite weak evidence on surrogacy of PFS for OS [7, 19], it could be speculated that the agencies may expect a PFS gain to be relevant to patients [17]. Considerations that may be relevant are the size of the PFS gain, the indication and stage of disease, and existing treatments or other supporting evidence. For example, evidence suggests that granting access for lung cancer drugs that prolong PFS by more than 3 months is robustly beneficial [15]. But the researchers also stress that this is likely to vary considerably among indications. We think that reporting the considerations about each end point, and explicitly stating whether PFS is seen as a surrogate or patient-relevant end point in the HTA reports, as in German reports, would increase transparency and facilitate harmonization.

In addition, recent initiatives by clinicians to define clinical relevance [6, 20] are a step forward. The European and American society for oncology have independently standardized approaches to grade the net health benefit, taking into account the clinical and safety results of medicines, compared with available treatments [6, 20]. This seems to be an important step toward consistent, transparent, and informed decision-making in a field of rapid development such as that of oncology treatments.

This study shows that the consideration of end point data varies between HTA jurisdictions. Further divergences are also seen between HTA bodies and drug regulatory agencies [5] because the regulator is willing to accept a higher degree of clinical uncertainty to expedite access to therapies. Currently, the development of anticancer drugs is designed to meet drug licensing requirements, and do not specifically accommodate the

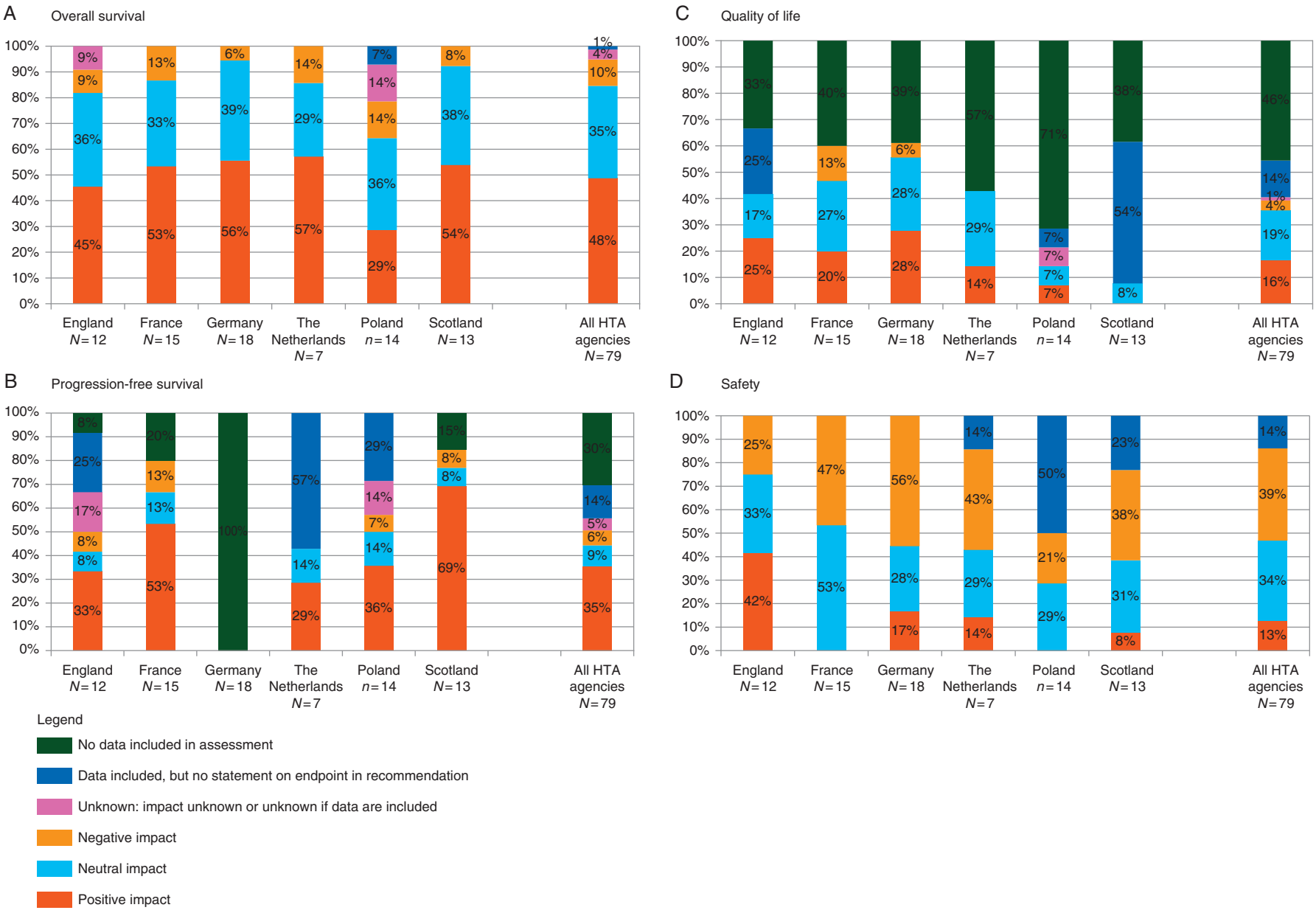


Figure 3. The impact of the end points on the recommendations: (A) overall survival, (B) progression-free survival, (C) quality of life, and (D) safety.

doi:10.1093/annonc/mdw233 | 7

requirements of HTA. A multi-stakeholder debate would be essential to align concrete robust evidence requirements in oncology and standardize the definition of 'relevant clinical benefit', which will benefit patients and society in general.

limitations

This study has some limitations. First, this study's results simplify real-world decision-making. We focused on REAs, but other factors such as cost-effectiveness (e.g. Scotland) can influence the recommendations. Further research would help to understand the impact of these end points on cost-effectiveness analysis, which was beyond the scope of this study. Moreover, our research is based on publicly available information, but other factors that are not reported may have had an influence in these complex decision-making processes. Secondly, we compared a limited number of HTA jurisdictions, although this is mitigated by their diversity, as we included both jurisdictions where cost-effectiveness is and is not relevant. Thirdly, interpreting value statements in the HTA reports is subjective. To standardize the interpretation, we introduced a decision algorithm with a quality control procedure, and consulted HTA experts to reduce possible misinterpretation.

conclusions

European decision-making on relative effectiveness of anticancer medicines is affected by a gap in requested clinical evidence versus the evidence that is actually available. OS and QoL are relevant to patients, but conclusive data on these end points are not always available (e.g. immature OS or incomplete/no QoL data), mainly because the regulator is willing to accept some degree of clinical uncertainty. At the same time, HTA agencies perceive the relevance of PFS differently. A multi-stakeholder debate would be essential to align concrete robust evidence requirements in oncology and a collectively shared definition for relevant clinical benefit, which will benefit patients and society in general.

acknowledgements

We would like to thank Dr Ilse Verstijnen for her help with the development and validation of the data abstraction form.

funding

This work was funded by the Zorginstituut Nederland (Dutch National Health Care Institute). No grant number is applicable.

disclosure

The authors have declared no conflicts of interest.

references

1. Sorensen S. Valuing end-of-life care in the United States: the case of new cancer drugs. Health Econ Pol Law 2012; 7: 411–430.
2. Sorenson C. Use of comparative effectiveness research in drug coverage and pricing decisions: a six-country comparison. Issue Brief (Commonw Fund) 2010; 91: 1–14.
3. Kleijnen S, George E, Goulden S et al. Relative effectiveness assessment of pharmaceuticals: similarities and differences in 29 jurisdictions. Value Health 2012; 15: 954–960.
4. Kleijnen S, Fathallah M, Van der Linden MW et al. Can a joint assessment provide relevant information for national/local relative effectiveness assessments? An in-depth comparison of pazopanib assessments. Value Health 2015; 18: 663–672.
5. Shah KK, Mestre-Ferrandiz J, Towse A, Smyth EN. A review of health technology appraisals: case studies in oncology. Int J Technol Assess Health Care 2013; 29: 101–109.
6. Cherny NI, Sullivan R, Dafni U 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: 1547–1573.
7. 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: 1389–1398.
8. Robinson AG, Booth CM, Eisenhauer EA. Progression-free survival as an end-point in solid tumours—perspectives from clinical trials and clinical practice. Eur J Cancer 2014; 50: 2303–2308.
9. Booth CM, Eisenhauer EA. Progression-free survival: meaningful or simply measurable? J Clin Oncol 2012; 30: 1030–1033.
10. Light DW, Lexchin J. Why do cancer drugs get such an easy ride? BMJ 2015; 350: h2068.
11. Sobrero AF, Pastorino A, Sargent DJ, Bruzzi P. Raising the bar for antineoplastic agents: how to choose threshold values for superiority trials in advanced solid tumors. Clin Cancer Res 2015; 21: 1036–1043.
12. van Harten WH, Wind A, de Paoli P et al. Actual costs of cancer drugs in 15 European countries. Lancet Oncol 2016; 17(1): 18–20.
13. Lim CS, Lee YG, Koh Y, Heo DS. International comparison of the factors influencing reimbursement of targeted anti-cancer drugs. BMC Health Serv Res 2014; 14: 595.
14. Mihajlović J, Dolk C, Tolley K et al. Reimbursement of targeted cancer therapies within 3 different European health care systems. Clin Ther 2015; 37(2): 474–480.
15. Lakdawalla DN, Chou JW, Linthicum MT et al. Evaluating expected costs and benefits of granting access to new treatments on the basis of progression-free survival in non-small-cell lung cancer. JAMA Oncol 2015; 1(2): 196–202.
16. Gemeinsamer Bundesausschuss. Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Afatinib. 2015. Berlin, Germany. https://www.g-ba.de/downloads/40-268-3424/2015-11-05_AM-TL-XII_Afatinib_2015-05-15-D-163_TrG.pdf (7 December 2015, date last accessed).
17. Chabot I, Rocchi A. Oncology drug health technology assessment recommendations: Canadian versus UK experiences. Clinicoecon Outcomes Res 2014; 6: 357–367.
18. EUnetHTA. Endpoints used in relative effectiveness assessment of pharmaceuticals: Surrogate Endpoints. http://www.eunethta.eu/outputs/methodological-guideline-rea-pharmaceuticals-surrogate-endpoints (7 December 2015, date last accessed).
19. Institute for Quality and Efficiency in Healthcare. Validity of surrogate endpoints in oncology: executive summary. 2011.
20. Ellis LM, Bernstein DS, Voest EE et al. American Society of Clinical Oncology perspective: raising the bar for clinical trials by defining clinically meaningful outcomes. J Clin Oncol 2014; 32: 1277–1280.