VALUE IN HEALTH 21 (2018) 698–706

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Effect of Crossover in Oncology Clinical Trials on Evidence Levels in Early Benefit Assessment in Germany

Georg Isbary, MD[1], Thomas R. Staab, MSc[1,2], Volker E. Amelung, PhD[2], Charalabos-Markos Dintsios, MPH[3], Christof Iking-Konert, MD[4], Sonja Mariotti Nesurini, PhD[5], Miriam Walter, PhD[5], Jörg Ruof, MD, MPH, MBA[2,6,,†]*

[1]*Roche Pharma AG, Grenzach-Wyhlen, Germany;* [2]*Medical School of Hanover, Hanover, Germany;* [3]*Health Services Research and Health Economics, Heinrich Heine University, Düsseldorf, Germany;* [4]*University Clinic Hamburg-Eppendorf, Hamburg, Germany;* [5]*nspm ltd, Meggen, Switzerland;* [6]*r-connect ltd, Basel, Switzerland*

ABSTRACT

Background: In oncology clinical trials, crossover is used frequently but may lead to uncertainties regarding treatment effects. **Objective:** To investigate the handling of evidence from crossover trials by the European Medicines Agency (EMA) and the German Federal Joint Committee (G-BA). **Methods:** For oncology medicines with early benefit assessments before January 2015, presence of crossover, clinical data, EMA requests for additional data, and G-BA benefit ratings/evidence levels were analyzed from manufacturers' dossiers, G-BA appraisals, European Public Assessment Reports, and original publications. **Results:** Eleven of 21 benefit assessments included crossover trials. Significant intergroup differences ($P < 0.05$) in overall survival (OS) were noted in 7 of 11 trials with and 7 of 10 without crossover. For 6 of 11 medicines with crossover, these were demonstrated before crossover. Treatment effects generally worsened with increasing proportions of crossover. The EMA requested additional data more frequently if crossover was performed, particularly if no OS data were available before crossover. The G-BA granted a considerable benefit to 73% of medicines with crossover and 40% of those without. Evidence levels were intermediate for 50% and 75%, respectively. None of the medicines received the highest evidence level. **Conclusions:** In G-BA appraisals, oncology medicines with crossover received better additional benefit ratings, but were assigned lower evidence levels, than those without. The five medicines with crossover after progression were assigned lower evidence levels than the six medicines with crossover after demonstration of superior OS, indicating that the way in which crossover is implemented may be one factor influencing the assignment of evidence levels by the G-BA.

Keywords: AMNOG, crossover design, early benefit assessment, evidence level, health technology assessment, oncology.

Introduction

In oncology clinical trials, "crossover" usually describes trial designs that allow patients from the control group to switch to the intervention arm and receive the investigational product after a predefined study event, for example, either after disease progression or after demonstration of clinical superiority of the investigational medicine [1]. Such designs are used frequently to maximize the number of patients who have access to the investigational drug in the study [2,3]. Moreover, they facilitate recruitment by increasing the trial's attractiveness to candidate patients [2,3] because patients may be more willing to enroll in a trial in which they are guaranteed to receive a given experimental treatment at some point. This is particularly true when early-phase data suggest a substantial treatment effect of the investigational agent. However, implementing crossover in a trial will reduce the treatment differences between the randomized arms for long-term trial end points, such as overall survival (OS) [1,2,4] and thus can influence a study's ability to answer a clinical question [5]. Results from a simulation study indicate that a crossover rate of more than 50% dramatically decreases the probability of detecting differences in OS by up to 90% [6].

Favorable effects on OS are still considered the most persuasive outcomes of a clinical trial in oncology [4]. The European Medicines Agency (EMA) recommends that crossover after disease progression should generally be avoided and be used only if there is confidence that the objectives of the trial can be met and adequate conclusions can be drawn [4,7].

Given the challenges in determining OS, particularly in the presence of crossover, additional efficacy end points, such as progression-free survival (PFS), are frequently used in clinical trials in oncology. There is an ongoing debate whether PFS is a useful surrogate for OS in oncology [8,9]. Although PFS may be

GI and TRS contributed equally to this work.

* *Address correspondence to:* Jörg Ruof, r-connect, Hauensteinstrasse 132, 4059 Basel, Switzerland.
E-mail: joerg.ruof@bluewin.ch
†Current address: r-connect, Basel, Switzerland.

1098-3015$36.00 – see front matter

http://dx.doi.org/10.1016/j.jval.2017.09.010

accepted as a primary end point in oncology trials by regulatory agencies in certain settings and under certain conditions [7], reliable and unbiased OS estimates are still likely to be required for health technology assessments [10].

In Germany, an early benefit assessment against the appropriate comparator therapy has been mandatory for new medicines since January 2011 [11–13]. The German Federal Joint Committee (*Gemeinsamer Bundesausschuss* [G-BA]) as the decision body and the Institute for Quality and Efficiency in Health Care (*Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen* [IQWiG]) as the scientific assessment body are charged with evaluating a medicine's additional benefit. This evaluation is based on a dossier submitted by the manufacturer and is conducted according to the principles of evidence-based medicine [11,14,15]. In addition to the extent of additional benefit versus the appropriate comparator therapy—major, considerable, minor, not quantifiable, or no additional benefit (or less benefit) [13,14]—the quality of the evidence base is evaluated. The evidence level is rated as proof (A), indication (B), or hint (C) on the basis of the number and characteristics of the submitted studies, the certainty of the results, and the consistency of the observed treatment effects [15]. The highest evidence level (A) requires a statistically significant effect in a meta-analysis or at least two independent randomized controlled trials showing statistically significant treatment effects in the same direction. Lower evidence levels are assigned when the presented evidence is based on only one randomized controlled trial or is considered to have a higher potential for bias [15]. Both the extent of additional benefit and the evidence level, together with other parameters, are considered by the National Association of Statutory Health Insurance Funds (*Spitzenverband Bund der Krankenkassen*) during the subsequent price negotiations with the manufacturer [16].

Currently, limited experience exists regarding the view of health technology assessment agencies on crossover trial design (e.g., the impact of crossover on the perceived evidence quality) [17]. We investigated the impact of crossover trials on the evaluations by the EMA and by the G-BA. Specifically, we determined whether the G-BA considered the way crossover was used in its appraisal and whether crossover impacted the evidence levels granted.

Methods

Oncology medicines with benefit assessments completed before January 1, 2015, were evaluated [18]. Orphan drugs were excluded from the analysis because an additional benefit is granted to such drugs by law [12]. Medicines for which no dossier was submitted were excluded because the basis for their assessment was missing.

For all medicines included in the analysis, the assessed indication was recorded. Where a re-assessment of the same medicine for the same indication had been performed, only the most recent assessment was analyzed.

Trials considered in G-BA appraisals were regarded as relevant for the analyses. Manufacturers' dossiers and G-BA appraisals, both obtained from the G-BA Web site [18], European Public Assessment Reports [19], and original trial publications were used as source documents. The analyses of the selected medicines encompassed the following points:

1. Presence of crossover trials
2. Clinical data
3. EMA's assessment
4. Benefit ratings and evidence levels granted by the G-BA

Presence of Crossover Trials

For this analysis, crossover was defined as a switch from a treatment specified for one group to a treatment specified for another group or to another treatment not included in the trial protocol that could occur either after disease progression or after reaching a predefined efficacy threshold. Medicines included in this analysis were divided into medicines with and without crossover studies.

Clinical Data

For all studies included in the analysis, the comparator therapy and the primary end point were extracted from the G-BA appraisal or the manufacturer's dossier. For studies involving crossover, the manufacturer's dossier was searched for a justification of the crossover design. Data on OS, PFS, and other primary end points were evaluated. OS and PFS were chosen because they are the most commonly used primary end points in oncology trials [20]. For studies with more than one data cut, data from the cut mentioned in the G-BA appraisal or, if none was mentioned, the most recent cut, were used for evaluation. If the G-BA analyzed several patient subgroups, only data for the subgroup with the best benefit rating were considered. Statistically significant differences ($P < 0.05$) between intervention and control treatment were assessed.

For medicines with crossover studies, it was determined whether data were available before crossover and whether there were statistically significant differences ($P < 0.05$) in OS, PFS, or other primary end points between the treatment arms before crossover.

Medicines with crossover in at least one clinical trial that presented OS or PFS data from different data cuts and provided information on the proportions of crossover patients at these time points were identified to assess the development of treatment effects over time in crossover trials.

EMA's Assessment

A conditional marketing authorization may be granted by the EMA in the absence of comprehensive clinical data to ensure immediate access to the medicine also outside of clinical trials [21]. In addition, the EMA may request postauthorization measures to obtain additional data and enable a more accurate assessment of the safety or efficacy of a new medicine [22].

Requests for additional data by the EMA (conditional marketing authorizations or postauthorization measures concerning OS) as stated in the European Public Assessment Reports were identified.

Benefit Ratings and Evidence Levels Granted by the G-BA

Overall additional benefit ratings assigned to the analyzed medicines were extracted from G-BA appraisals. The additional benefit rating for mortality was analyzed separately, because mortality (i.e., OS) is the preferred benefit category in assessments of oncology medicines [20] and because OS data can be strongly influenced by crossover [2,3,7,10,17]. Whether safety data had a positive, negative, or neutral effect on the overall benefit rating was also assessed. The level of evidence granted by the G-BA in its appraisals was analyzed for medicines with positive additional benefit ratings. For medicines for which the G-BA analyzed more than one patient subgroup, only the benefit rating and evidence level for the subgroup with the highest additional benefit were taken into account.

To assess the impact of crossover on benefit assessments, a comparison of benefit ratings and evidence levels was performed for medicines with and without crossover studies.

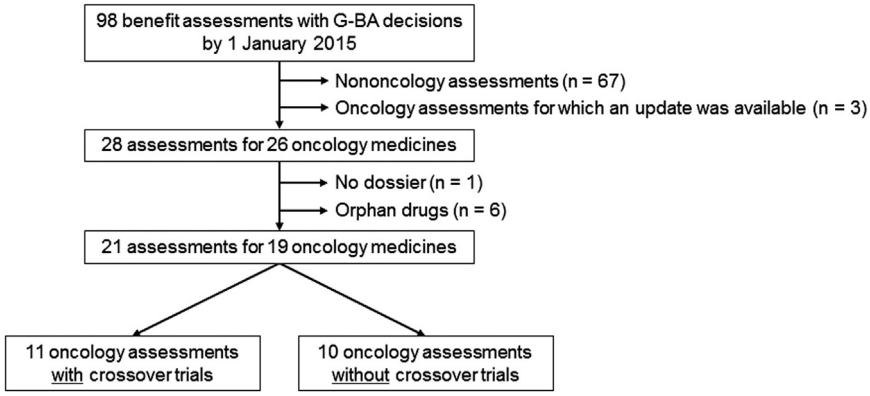

Fig. 1 – Benefit assessments for oncology medicines with and without crossover in clinical trials (analysis set). G-BA, *Gemeinsamer Bundesausschuss* (Federal Joint Committee).

Results

By January 1, 2015, 26 oncology medicines had completed 28 benefit assessments. Seven assessments (six orphan drugs and one medicine for which no dossier was submitted) were excluded from the analysis. Therefore, 21 assessments of 19 medicines form the basis of the presented analyses (Fig. 1). Table 1 presents the indication, the comparator therapy, and the primary end point used in the relevant studies.

Presence of Crossover Trials

For 11 of 21 assessments (52%), at least 1 trial assessed by the G-BA included crossover (Fig. 1). Information on why a crossover design was used is presented in Table 1.

Clinical Data

Table 2 summarizes the data reported for the medicines with and without crossover studies.

All 11 medicines with crossover studies reported OS data (median OS and/or hazard ratio). In addition, 10 medicines (91%) presented data on PFS or a primary end point other than OS/PFS. Significant differences ($P < 0.05$) versus the control treatment were presented for 7 of 11 medicines for OS (64%) and for all 10 medicines presenting such data for PFS or a primary end point other than OS/PFS. Six of the seven medicines with a significant difference ($P < 0.05$) in OS between treatment arms demonstrated a positive effect before crossover, whereas for all four medicines without significant differences in OS ($P \geq 0.05$), no data were available before crossover (Table 2).

OS data were reported in 9 of 10 benefit assessments without crossover studies (90%). The remaining assessment was based on a single-arm study and presented mortality data in terms of the 1-year rate of OS. PFS was presented in 7 of 10 assessments (70%). Two of these additionally reported data on a primary end point other than OS/PFS. Significant treatment effects ($P < 0.05$) on OS were demonstrated in 7 of 10 assessments (70%). Six of 10 assessments (60%) presented significant differences ($P < 0.05$) in PFS or primary end points other than OS/PFS.

OS data from different data cuts with varying proportions of crossover patients were available for 5 of 11 medicines with crossover trials (Fig. 2A). Similar data for PFS were provided for 4 of 11 medicines (Fig. 2B). For three of these medicines, data before and after crossover were available. For the remaining medicines, only time points after crossover were reported. The hazard ratios for OS and PFS generally increased (i.e., the treatment effect decreased) with an increasing proportion of patients who switched treatment. For ruxolitinib, the hazard ratios for OS were almost unaffected by the proportion of crossover patients.

EMA's Assessment

Five of the 11 oncology medicines with crossover (46%) had a postauthorization measure related to OS (n = 4) and/or held a conditional marketing authorization (n = 2) (Table 2). For most of these medicines (three of five), no OS data were reported before crossover (Table 2). Only 2 of 10 medicines without crossover (20%) had a conditional marketing authorization, and none of them were subject to an OS-related postauthorization measure.

Benefit Ratings and Evidence Levels Granted by the G-BA

G-BA benefit ratings, in terms of both OS and overall, and evidence levels are listed in Table 2 and summarized in Figure 3. The influence of safety data on the assessment of overall benefit is presented in Table 2.

In terms of mortality, the G-BA granted an additional benefit to 8 of 11 (73%) medicines with crossover and to 6 of 10 (60%) medicines without crossover (Fig. 3A). Eight of 11 medicines with crossover (73%), but only 4 of 10 without crossover (40%), were granted a considerable overall additional benefit (Fig. 3B).

Evidence levels were evaluated for 10 and 8 medicines with and without crossover that had received positive additional benefit ratings. None of the medicines were granted the highest evidence level (A). For half of the medicines with crossover (5 of 10 [50%]) and 6 of 8 medicines without crossover (75%), the evidence level was considered to be intermediate (B). In summary, benefit ratings were better for medicines with crossover studies than for those without. However, the granted evidence levels were lower.

Discussion

In clinical trials in oncology, particularly when OS is the primary end point, crossover trial designs that allow patients randomized to the control arm to cross over to the investigational treatment play an important role. However, the crossover design limits the robustness of the evidence generated, and finding the right balance between fulfilling scientific evidence requirements and ensuring timely patient access to a superior treatment, both in the study and outside of clinical trials, is a sensitive issue. This is especially important in light of the increasing influence of payers and their high standards regarding evidence generation. From our point of view, two types of trial designs have to be taken into consideration: 1) crossover after disease progression is permitted

Table 1 – Indication, comparator therapy, primary end point, and reason for crossover (if applicable) in studies submitted by the manufacturer for oncology medicines with and without crossover studies.

Medicine	Indication	Comparator therapy	Primary end point	Reason for crossover
Oncology medicines with crossover studies				
Abiraterone acetate (second indication)	Prostate cancer (after failure of androgen withdrawal)	Placebo[*]	OS and PFS[†,‡]	Primary end point reached in interim analysis[§]
Afatinib	NSCLC	Cisplatin and pemetrexed	PFS[‡]	After progression or intolerable toxicity (crossover ethically required if targeted therapies show early efficacy)
Crizotinib	NSCLC	Pemetrexed or docetaxel	PFS[‡]	After progression (desire for unbiased OS assessment has to be weighed against ethical demands)
Dabrafenib	Melanoma	Dacarbazine	PFS[‡]	After progression (at the time of study design, phase I data suggested marked superiority of investigational product over the planned comparator)
Pertuzumab	Breast cancer	Placebo[‖]	PFS[‡]	Primary end point reached in interim analysis
Radium-223-dichloride	Prostate cancer	Placebo[¶]	OS	Primary end point reached in interim analysis
Regorafenib	Metastatic colorectal cancer	Placebo[¶]	OS	Primary end point reached in interim analysis
Ruxolitinib	Chronic myeloproliferative disorders	Placebo	Proportion of subjects achieving ≥35% reduction in spleen volume	After progression (OS was not primary analysis goal; therefore, crossover permitted for ethical reasons)
Trastuzumab emtansine	Breast cancer	Lapatinib and capecitabine	OS and PFS[†,‡]	Primary end point reached in interim analysis
Vandetanib	Thyroid neoplasms	Placebo	PFS[‡]	After progression (no other approved therapy available at the time of study)
Vemurafenib	Melanoma	Dacarbazine	OS and PFS[†,‡]	Primary end point reached in interim analysis
Oncology medicines without crossover studies				
Abiraterone acetate (first indication)	Prostate cancer (after chemotherapy)	Placebo[#]	OS	NA
Aflibercept	Colorectal cancer	Placebo[**]	OS	NA
Axitinib	Renal cell carcinoma	Sorafenib	PFS[‡]	NA
Cabazitaxel	Prostate cancer	Mitoxantrone	OS	NA
Enzalutamide	Prostate cancer	Placebo[¶]	OS	NA
Eribulin	Breast cancer	Treatment of physician's choice (EMBRACE study) Capecitabine (study 301)	OS (EMBRACE study) OS and PFS[†,‡] (study 301)	NA
Ipilimumab (first indication)	Melanoma (second line)	Placebo[¶]	OS	NA
Ipilimumab (second indication)	Melanoma (first line)	Various[††]	Various[††]	NA
Pixantrone	Non-Hodgkin lymphoma	Monotherapy with one of seven agents permitted per protocol (to be chosen by the physician)	CR/Cru[‡]	NA
Vismodegib	Basal cell carcinoma	Single-arm study	ORR[‡]	NA

CR/CRu, complete response and unconfirmed complete response; EMBRACE, Eisai Metastatic Breast Cancer Study Assessing Physician's Choice versus E7389; NA, not applicable; NSCLC, non–small cell lung cancer; ORR, overall response rate; OS, overall survival; PFS, progression-free survival.

[*] Both arms plus prednisone/prednisolone.

[†] OS and PFS were coprimary end points.

[‡] PFS, CR/Cru, and ORR were composite primary end points.

[§] The prespecified stopping boundary ($P < 0.001$) was reached only for PFS but crossover was recommended by the data safety monitoring board.

[‖] Both arms plus trastuzumab emtansine and docetaxel.

[¶] Both arms plus best supportive care.

[#] Both arms plus prednisone.

[**] Both arms plus irinotecan, 5-fluorouracil, and folinic acid (FOLFIRI).

[††] The manufacturer presented a pooled analysis of several studies for a head-to-historic-head comparison because no study against the appropriate comparator therapy as defined by the G-BA was available.

Table 2 – Data on OS, PFS, or primary end points other than OS/PFS, EMA requests for OS data, G-BA benefit ratings, and G-BA evidence level.

Oncology medicines	OS,* absolute values† HR (95% CI)	PFS or primary end points other than OS/PFS,* absolute values† HR (95% CI)	EMA request for OS data	G-BA benefit rating in terms of mortality*,‡	Impact of safety data on overall G-BA benefit rating*,‡	Overall G-BA benefit rating*,‡	Overall G-BA evidence level*,‡
With crossover studies							
Abiraterone acetate§ (second indication)	**35.3 vs 30.1 0.79 (0.66–0.96)**	**16.5 vs 8.3 0.53 (0.45–0.62)**	No	Considerable	Neutral	2	B
Afatinib	**31.6 vs 21.1 0.55 (0.36–0.85)**	**13.7 vs 5.6 0.28 (0.18–0.44)**	No	Considerable	Neutral	2	B
Crizotinib	20.3 vs 22.8 1.02 (0.68–1.54)	**7.7 vs 3.0 0.49 (0.37–0.64)**	Conditional approval, PAM related to OS	No additional benefit	Neutral	2‖	C
Dabrafenib	18.2 vs 15.6 0.76 (0.48–1.21) *(indirect comparison)*	**5.1 vs 2.7 0.30 (0.18–0.51)** *(indirect comparison)*	No	No suitable evidence¶	No suitable evidence¶	5	–
Pertuzumab§	**NR 0.57 (0.44–0.74)**	**NR 0.55 (0.45–0.68)**	No	Considerable	Neutral	2	C
Radium-223-dichloride§	**14.9 vs 11.3 0.70 (0.58–0.83)**	ND	No	Considerable	Neutral	2	B
Regorafenib§	**6.5 vs 5.0 0.77 (0.64–0.94)**	**2.0 vs 1.7 0.49 (0.42–0.58)**	No	Relevant increase in OS	Negative	3	C
Ruxolitinib	Deaths (%): 27.1 vs 35.1 0.69 (0.46–1.03)	**Spleen reduction (%): 41.9 vs 0.7#**	PAM related to OS	Advantage in terms of OS**	Neutral	2	C
Trastuzumab emtansine§	**30.9 vs 23.7 0.70 (0.53–0.85)**	**9.0 vs 6.9 0.69 (0.55–0.85)**	PAM related to OS	Considerable	Positive	2	B
Vandetanib	Deaths (%): 16.7 vs 16.7 1.06 (0.26–4.38)	**Progression (%): 36.5 vs 58.3 0.47 (0.29–0.77)**	Conditional approval	No additional benefit	Negative	3‖	C
Vemurafenib§	**9.2 vs 7.8 0.37 (0.26–0.55)**	**5.3 vs 1.6 0.26 (0.20–0.33)**	PAM related to OS	Considerable	Neutral	2	B
Without crossover studies							
Abiraterone acetate (first indication)	**14.8 vs 10.9 0.65 (0.54–0.77)**	**5.6 vs 3.6 0.67 (0.59–0.78)**	No	Considerable	Neutral	2	B
Aflibercept	**13.5 vs 12.1 0.82 (0.71–0.94)**	**6.9 vs 4.7 0.76 (0.66–0.87)**	No	Relevant increase in OS	Negative	3	B
Axitinib	29.4 vs 27.8 0.81 (0.55–1.19)	**12.1 vs 6.5 0.46 (0.32–0.68)**	No	No additional benefit	Positive	3	B
Cabazitaxel	**15.1 vs 12.7 0.70 (0.59–0.83)**	**2.8 vs 1.4 0.74 (0.64–0.86)**	No	Considerable	Negative	3	B
Enzalutamide	**18.4 vs 13.6 0.63 (0.53–0.75) < 0.0001**	**8.3 vs 2.9 0.40 (0.35–0.47) < 0.0001**	No	Considerable	Positive	2	B
Eribulin	**16.0 vs 13.5 0.81 (0.71–0.92)** *(meta-analysis)*	NR	No	Considerable	Neutral	2	C
Ipilimumab (first indication)	**10.1 vs 6.4 0.66 (0.51–0.87)**	ND	No	Considerable	Neutral	2	B
Ipilimumab (second indication)	**NR 0.48 (0.37–0.64)** *(head-to-historic-head comparison)*	ND	No	No suitable evidence¶	No suitable evidence¶	5	–
Pixantrone	13.9 vs 7.8 0.76 (0.47–1.24)	**PFS (mo): 5.8 vs 2.8 0.50 (0.32–0.78) CR/CRu (%): 28.0 vs 4.1#**	Conditional approval	No suitable evidence††	No suitable evidence††	5	–
Vismodegib	1-y rate of OS (%): 91.6 (95% CI 83.5–99.7)‡‡	PFS (mo): 9.5 (95% CI 7.4–11.9)‡‡ ORR (%): 42.9 (95% CI 30.5–56.0)‡‡	Conditional approval	No suitable evidence	Negative	3‖	C

Note: Bold denotes significant differences between treatment arms (defined as $P < 0.05$).

CI, confidence interval; CR/CRu, complete response and unconfirmed complete response; EMA, European Medicines Agency; G-BA, *Gemeinsamer Bundesausschuss* (Federal Joint Committee); HR, hazard ratio; ND, not determined; NR, not reported; ORR, overall response rate; OS, overall survival; PAM, postauthorization measure; PFS, progression-free survival.

* In the case of subgroup analysis for a medicine, data for the subgroup with the highest additional benefit are shown.

† Data are median OS/PFS (mo), unless otherwise stated.

‡ In the case of a re-assessment, the most recent assessment is shown. Categories for the overall benefit rating: 1, major; 2, considerable; 3, minor; 4, not quantifiable; 5, no additional benefit; and 6, less benefit. Categories for the evidence level: A, proof; B, indication; C, hint.

§ Data available before crossover (applicable only for oncology medicines with crossover studies).

‖ Because of positive treatment effects on morbidity (vandetanib, vismodegib) or morbidity and quality of life (crizotinib). ¶Indirect/head-to-historic head comparison not accepted by the G-BA.

# $P < 0.05$.

** Based on additional post hoc interim analyses that were considered by the G-BA, although they were not presented in the dossier, and that showed a statistically significant difference in OS in favor of ruxolitinib.

†† Comparator therapy used in the study not accepted by the G-BA.

‡‡ Single-arm study; HR calculation not possible.

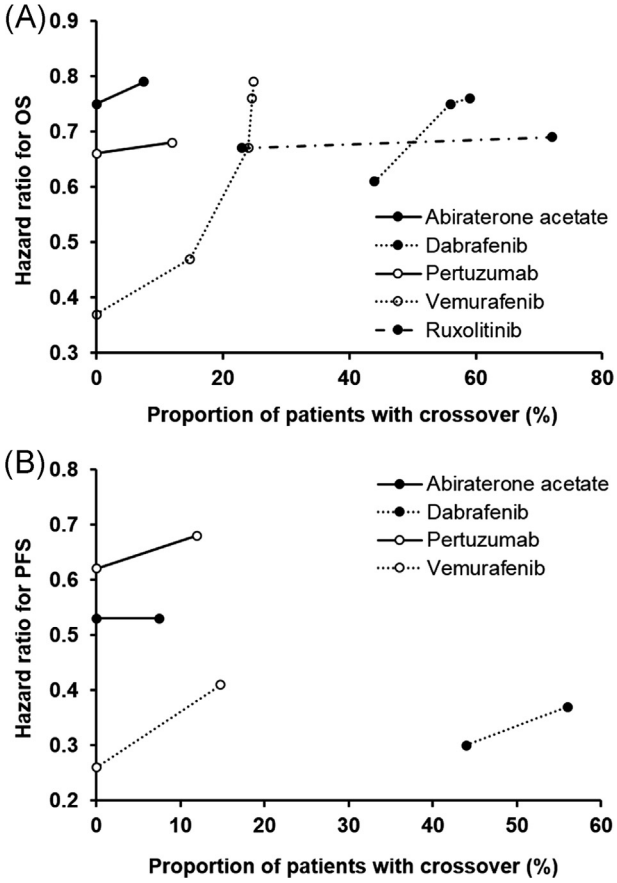

Fig. 2 – Impact of crossover on treatment effect. Hazard ratios for (A) OS and (B) PFS are shown in relation to the proportion of patients who switched therapy. OS, overall survival; PFS, progression-free survival.

from study start and 2) crossover occurs after a prespecified stopping boundary has been exceeded.

1. Crossover after disease progression from study start

   In the first scenario, the primary end point is typically PFS because no appropriate, unbiased OS data are generated before crossover. Crossover to the intervention arm is usually permitted when a patient in the control arm experiences disease progression [23]. In this case, all available OS data will be affected by crossover, and the true treatment effect will be difficult to discern. In many cases, a benefit in terms of disease progression will be observed, but there will be no convincing differences in OS. This can be observed for four of the five medicines with crossover after progression from study start in the present analysis (ruxolitinib, crizotinib, dabrafenib, and vandetanib; see Table 2). There are three possible explanations for these observations: the investigational medicine confers an OS benefit that is masked by crossover; there is a treatment effect on PFS but not on OS; or a decrease in cancer-related mortality is counterbalanced by an increase in non–cancer-related deaths [3].

   For some medicines with a substantial treatment effect on OS, it may be possible to demonstrate an OS benefit even in the presence of crossover [9]. For the fifth medicine with crossover in this analysis, afatinib, a statistically significant treatment effect on OS ($P < 0.05$) was demonstrated in one patient subgroup, in spite of crossover after progression throughout the study (see Table 2). However, the exact extent of the OS benefit is unknown, which will lead to uncertainty in subsequent evaluations [3].

   In certain types of cancers, OS is not considered the primary analysis goal, for example, in diseases with a high cure rate or a slow chronic course that would require extensive study periods or large patient numbers to allow meaningful OS evaluations [24,25]. In such cases, it may be considered necessary to permit crossover after progression from study start for ethical reasons: this was done for ruxolitinib in chronic myeloproliferative disorders [18]. However, the EMA stipulates that crossover after progression should be used only if an unfavorable effect on OS can be excluded [4], which may or may not be possible on the basis of the data obtained.

   The inherent uncertainty about OS data from trials with crossover from study start is reflected in the evidence levels granted by the G-BA. Out of five medicines for which such a design was used, one was rated as having no additional benefit and thus no evidence level, one received an intermediate evidence level, and three received the lowest level. Of note, the EMA requested additional OS data for all these three medicines, further highlighting the uncertainty associated with such data. In line with the requirements stipulated in EMA guidelines [4,7], the IQWiG recommends that trials should present at least one reliable interim analysis before crossover is introduced [26]. Therefore, the granting of low evidence levels by the G-BA appears justified if no such data are presented.

   If no reliable OS data are available, medicines are essentially approved on the basis of improvements in surrogate end points such as PFS or overall response rate. However, experience has shown that the use of surrogate end points is not without issues [9,27]. The European Network for Health Technology Assessment has suggested that the use of PFS data without supporting OS data should be confined to the adjuvant setting [28]. A recent systematic review of trial-level meta-analyses found that 52% of correlations between surrogate end points and OS were weak [9]. It has been argued that, conversely, 48% of correlations are not perfect, but reasonably strong [27], and that the association between OS and PFS will never be definite, even within a given indication [28]. However, it may be challenging to establish a valid association in the first place if OS data from available trials are affected by crossover or any subsequent therapy after patients stop study treatment [29], which can be expected to be particularly relevant to studies in the front-line setting [30].

2. Crossover after demonstration of superior efficacy in terms of OS

   The second scenario of a crossover trial design in oncology is the implementation of crossover following the demonstration that the investigational agent improves outcomes. This decision may be based on the crossing of a predefined stopping boundary in an interim analysis and/or endorsed by an independent monitoring board. In addition to the introduction of crossover, enrollment may be stopped [31]. Consequently, OS data unaffected by crossover will be available up to, but not beyond, the time of the interim analysis.

   Crossover was implemented following the demonstration of superior efficacy in terms of OS for six medicines included in this analysis. For four of these (pertuzumab, radium-223-dichloride, regorafenib, and trastuzumab emtansine), crossover had been implemented only after the most recent data cut. Consequently, only the long-term follow-up for OS, but none of the OS data presented in the dossier, was impacted by crossover. For two medicines (abiraterone acetate and vemurafenib), OS data before and after crossover were available [18].

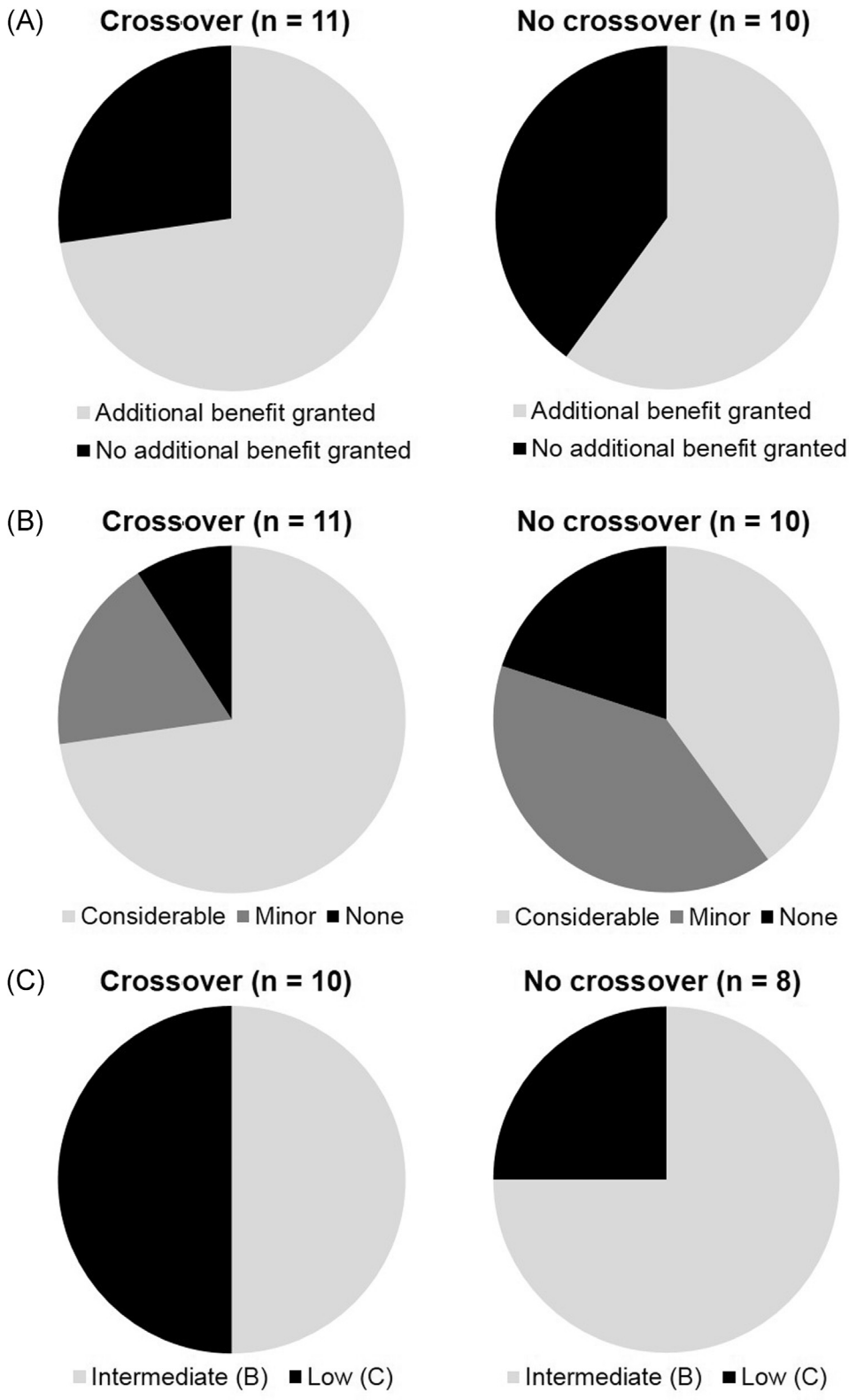

Fig. 3 – Additional benefit ratings—in terms of (A) mortality and (B) overall—and evidence levels* (C) granted by the G-BA. G-BA, *Gemeinsamer Bundesausschuss* (Federal Joint Committee). *Medicines that granted no additional benefit were excluded from the evaluation of evidence levels (one and two benefit assessments with and without crossover, respectively).

Especially in long-term trials, interim analyses are performed to monitor whether the available data on efficacy and/or safety indicate that the trial should be modified or stopped. This is also recognized by the EMA, which cautions that the risk of damaging the integrity of a trial must be minimized, particularly in late-stage trials [32]. Generally,

when considering stopping a trial early, obtaining the most accurate estimate for the treatment effect must be weighed against minimizing the number of trial participants who receive an inferior treatment [33]. Stopping trials early appears to have become more common in oncology [31] and in general [34]. Commercial motives have been suspected [31,35], but alternative explanations, such as the increasing use of targeted therapies and improved application of statistical principles, are also possible [36]. The main concerns associated with premature study termination are the fact that the differences observed may represent only a random high or low [33,37,38], and/or the large treatment effects observed in these trials, especially with low numbers of events [34,38]. However, such large effects can be expected when stopping rules are used, because a large outcome difference is required to generate a significance level exceeding a predefined level in an early interim analysis [39]. Therefore, the fact that a trial is stopped prematurely should not *per se* be interpreted to mean that the estimated effect size is incorrect [33]. Results from a simulation study confirmed that for trials with well-designed interim-monitoring plans, an inflation of the treatment effect is only likely for interim analyses based on 25% or less of the planned total events [40]. There is a general consensus that stopping a trial early may be acceptable if interim analyses are based on a sufficiently large number of events and use stopping rules associated with low *P* values [35,37]. Confirmatory analyses, for example, comprising also those patients not yet included in the interim analysis, may lend more credibility to the results [32,35]. Moreover, the clinical context should be considered [35,41].

For the six medicines with crossover after an interim analysis, the decision to implement crossover and/or stop the trial was made following interim analyses based on approximately 40% to 70% of the planned number of total deaths (18). *P* values were less than 0.01 for all medicines and less than 0.001 for three medicines. According to the simulation study cited above, an inflation of the treatment effect should therefore be unlikely. Data from additional analyses before crossover are mentioned in the dossiers of two medicines (radium-223-dichloride and regorafenib) and confirm the results of the interim analyses. For the other four medicines, additional OS analyses after crossover are available either in the dossiers (abiraterone acetate and vemurafenib [18]) or from the literature (pertuzumab [42] and trastuzumab emtansine [43]) and demonstrate sustained superiority in spite of crossover. Therefore, it can be concluded that the conduct of the studies in question meets high standards in all six cases. This is supported by the observation that the EMA requested additional OS data for only two of these medicines (trastuzumab emtansine and vemurafenib) and that these requests exclusively referred to already planned evaluations of OS to be performed after approval.

The G-BA accorded the lowest evidence level to two medicines, regorafenib and pertuzumab, because of an unclear relevance of the study to the German context and uncertainties regarding non-OS data. The other four of the six medicines, including those for which the EMA had requested additional data, received an intermediate evidence level. In two cases, this was due to data-related uncertainties. However, for the remaining two medicines (abiraterone acetate and radium-223-dichloride), the only reason why the highest evidence level was not granted was the conduct of only one study. The same argument was put forward for three of the six medicines without crossover that were granted an intermediate evidence level [18], suggesting that the conclusiveness of trials with crossover after a meaningful interim analysis may be considered similar to that of trials without crossover.

IQWiG guidance maintains that generally two pivotal trials are required to receive the highest evidence level [15]. Especially in oncology, however, ethical considerations often make the conduct of a second trial challenging if convincing data indicating superiority of the investigational treatment, particularly with regard to OS, have been obtained in a first trial. This is due to the general perception that uncertainty regarding which of the analyzed treatments is more effective is required for a randomized trial to be ethical [44].

A limitation of this analysis is the relatively small data set, because the number of available oncology assessments is limited. More differentiated findings regarding G-BA's view on crossover are expected to emerge in the coming years when more medicines have undergone the assessment procedure. Furthermore, this analysis is based on a conservative estimation because assessments for the best subgroup were considered in the case of subgroup analyses, irrespective of subgroup size. This is likely to have led to a positive shift in the G-BA ratings.

Conclusions

Oncology medicines with crossover trials were granted better additional benefit ratings by the G-BA than those without. However, the assigned evidence levels were lower. Within the group of medicines with crossover trials, the assigned evidence levels were lower for five medicines with crossover after progression from study start than for six medicines with crossover following demonstration of significant OS differences ($P < 0.05$). The latter were assigned evidence levels comparable to those of medicines with no crossover, indicating that the way in which crossover is implemented may be one factor influencing the evidence levels assigned by the G-BA. None of the medicines received the highest evidence level, regardless of crossover.

REFERENCES

[1] Ishak KJ, Proskorovsky I, Korytowsky B, et al. Methods for adjusting for bias due to crossover in oncology trials. Pharmacoeconomics 2014;32:533–46.

[2] Latimer NR, Abrams KR, Lambert PC, et al. Adjusting survival time estimates to account for treatment switching in randomized controlled trials–an economic evaluation context: methods, limitations, and recommendations. Med Decis Making 2014;34:387–402.

[3] Prasad V, Grady C. The misguided ethics of crossover trials. Contemp Clin Trials 2014;37:167–9.

[4] European Medicines Agency. Guideline on the evaluation of anticancer medicinal products in man. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/01/WC500137128.pdf. [Accessed January 14, 2016].

[5] Prasad V. Use of the word "cure" in the oncology literature. Am J Hosp Palliat Care 2014;32:477–83.

[6] Zhang L, Ko C-W, Tang S, et al. Relationship between progression-free survival and overall survival benefit—a simulation study. Ther Innov Regul Sci 2013;47:95–100.

[7] European Medicines Agency. Appendix 1 to the guideline on the evaluation of anticancer medicinal products in man—methodological consideration for using progression-free survival (PFS) or disease-free survival (DFS) in confirmatory trials. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/01/WC500137126.pdf. [Accessed January 14, 2016].

[8] 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;1:196–202.

[9] Prasad V, Kim C, Burotto M, et al. 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.

[10] Watkins C, Huang X, Latimer N, et al. Adjusting overall survival for treatment switches: commonly used methods and practical application. Pharm Stat 2013;12:348–57.

[11] Bundesministerium für Gesundheit. [AM-NutzenV]. Available from: http://www.gesetze-im-internet.de/bundesrecht/am-nutzenv/gesamt.pdf. [Accessed January 14, 2016].

[12] Deutscher Bundestag. [Sozialgesetzbuch (SGB) Fünftes Buch (V) - Gesetzliche Krankenversicherung - (Artikel 1 des Gesetzes v. 20. Dezember 1988, BGBl. I S. 2477, letzte Änderung durch Artikel 3 des Gesetzes vom 20. Dezember 2012 (BGBI. I S. 2781))]. Available from: http://www.gesetze-im-internet.de/bundesrecht/sgb_5/gesamt.pdf. [Accessed January 14, 2016].

[13] Ruof J, Schwartz FW, Schulenburg JM, et al. Early benefit assessment (EBA) in Germany: analysing decisions 18 months after introducing the new AMNOG legislation. Eur J Health Econ 2014;15:577–89.

[14] G-BA. Chapter 5: assessment of the benefits of pharmaceuticals according to §35a SGB V. Available from: http://www.english.g-ba.de/downloads/17-98-3042/2011-05-18_5%20Kapitel%20VerfO%20Englisch.pdf. [Accessed January 14, 2016].

[15] *Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen*. Allgemeine Methoden 4.2. Available from: https://www.iqwig.de/download/IQWiG_Methoden_Version_4-2.pdf. [Accessed January 14, 2016].

[16] GKV Spitzenverband. [AMNOG-Informationen - Thema: Zuzsatznutzen und Erstattungsbetrag]. Available from: https://www.gkv-spitzenverband.de/media/dokumente/krankenversicherung_1/arzneimittel/amnog_aktuell/20160114_AMNOG_Informationen_4_Zusammenhang_Zusatznutzen_und_EB_korr.pdf. [Accessed January 14, 2016].

[17] Jönsson L, Sandin R, Ekman M, et al. Analyzing overall survival in randomized controlled trials with crossover and implications for economic evaluation. Value Health 2014;17:707–13.

[18] G-BA. Overview of products [Übersicht der Wirkstoffe]. Available from: http://www.g-ba.de/informationen/nutzenbewertung/. [Accessed January 14, 2016].

[19] European Medicines Agency. European public assessment reports. Available from: http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/landing/epar_search.jsp&mid=WC0b01ac058001d124. [Accessed January 14, 2016].

[20] Ruof J, Knoerzer D, Duenne AA, et al. Analysis of endpoints used in marketing authorisations versus value assessments of oncology medicines in Germany. Health Policy 2014;118:242–54.

[21] European Medicines Agency. Commission Regulation (EC) No 507/2006 of 29 March 2006 on the conditional marketing authorisation for medicinal products for human use falling within the scope of Regulation (EC) No 726/2004 of the European Parliament and of the Council. Available from: http://ec.europa.eu/health/files/eudralex/vol-1/reg_2006_507/reg_2006_507_en.pdf. [Accessed January 14, 2016].

[22] European Medicines Agency. Post-authorisation measures: questions and answers. Available from: http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/q_and_a/q_and_a_detail_000037.jsp. [Accessed January 14, 2016].

[23] Morden JP, Lambert PC, Latimer N, et al. Assessing methods for dealing with treatment switching in randomised controlled trials: a simulation study. BMC Med Res Methodol 2011;11:4.

[24] Wörmann B, Freund M, Overkamp F, et al. Status of clinical trials in oncology–2013 and onwards [in German]. Onkologie 2013;36(Suppl. 2):3–8.

[25] Soria JC, Massard C, Le CT. Should progression-free survival be the primary measure of efficacy for advanced NSCLC therapy? Ann Oncol 2010;21:2324–32.

[26] Vervölgyi V, Beckmann L. Beurteilung von Treatment Switching in Bewertungen des Instituts. IQWiG im Dialog. Cologne, Germany: IQWiG, 2014.

[27] Jena AB, Zhang J, Lakdawalla DN. The trade-off between speed and safety in drug approvals [published online ahead of print September 29, 2016]. JAMA Oncol. http://dx.doi.org/10.1001/jamaoncol.2016.3337.

[28] European Network for Health Technology Assessment. Endpoints used for relative effectiveness assessment: clinical endpoints. Available from: https://eunethta.fedimbo.belgium.be/sites/5026.fedimbo.belgium.be/files/explorer/WP7-SG3-GL-clin_endpoints_amend2015.pdf. [Accessed January 14, 2016].

[29] Diaz J, Sternberg CN, Mehmud F, et al. Overall survival endpoint in oncology clinical trials: addressing the effect of crossover–the case of pazopanib in advanced renal cell carcinoma. Oncology 2016;90:119–26.

[30] Blumenthal GM, Karuri SW, Zhang H, et al. Overall response rate, progression-free survival, and overall survival with targeted and standard therapies in advanced non-small-cell lung cancer: US Food and Drug Administration trial-level and patient-level analyses. J Clin Oncol 2015;33:1008–14.

[31] Trotta F, Apolone G, Garattini S, et al. Stopping a trial early in oncology: for patients or for industry? Ann Oncol 2008;19:1347–53.

[32] Committee for Medicinal Products for Human Use. Reflection paper on methodological issues in confirmatory clinical trials planned with an adaptive design. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003616.pdf. [Accessed January 14, 2016].

[33] Goodman SN. Stopping trials for efficacy: an almost unbiased view. Clin Trials 2009;6:133–5.

[34] Montori VM, Devereaux PJ, Adhikari NK, et al. Randomized trials stopped early for benefit: a systematic review. JAMA 2005;294:2203–9.

[35] Mueller PS, Montori VM, Bassler D, et al. Ethical issues in stopping randomized trials early because of apparent benefit. Ann Intern Med 2007;146:878–81.

[36] Bradbury P, Meyer R, Pater J, et al. Stopping a trial early in oncology: for patients or for industry? Ann Oncol 2009;20:395–6.

[37] Heffner AC, Milbrandt EB, Venkataraman R. Trials stopped early for benefit? Not so fast! Crit Care 2007;11:305.

[38] Wilcox RA, Djulbegovic B, Moffitt HL, et al. Randomized trials in oncology stopped early for benefit. J Clin Oncol 2008;26:18–9.

[39] Pater JL, Goss P, Meyer R. Stopping trials for benefit can (sometimes) benefit patients. J Clin Oncol 2008;26:2787–8: author reply 2788.

[40] Freidlin B, Korn EL. Stopping clinical trials early for benefit: impact on estimation. Clin Trials 2009;6:119–25.

[41] Whitehead J, Cobo E. Stopping a trial early in oncology: for patients or for industry? Ann Oncol 2008;19:1512–3.

[42] Swain SM, Baselga J, Kim SB, et al. Pertuzumab, trastuzumab, and docetaxel in HER2-positive metastatic breast cancer. N Engl J Med 2015;372:724–34.

[43] Diéras V, Miles D, Verma S, et al. Trastuzumab emtansine versus capecitabine plus lapatinib in patients with previously treated HER2-positive advanced breast cancer (EMILIA): a descriptive analysis of final overall survival results from a randomised, open-label, phase 3 trial. Lancet Oncol 2017;18:732–42.

[44] Kurzrock R, Stewart DJ. Equipoise abandoned? Randomization and clinical trials. Ann Oncol 2013;24:2471–4.