



# Local Investigators Significantly Overestimate Overall Response Rates Compared to Blinded Independent Central Reviews in Uncontrolled Oncology Trials: A Comprehensive Review of the Literature

Cinzia Dello Russo<sup>1,2</sup> and Pierluigi Navarra<sup>1\*</sup>

<sup>1</sup>Section of Pharmacology, Department of Healthcare Surveillance and Bioethics, Fondazione Policlinico Universitario A. Gemelli IRCCS, Università Cattolica Del Sacro Cuore, Rome, Italy, <sup>2</sup>MRC Centre for Drug Safety Science and Wolfson Centre for Personalized Medicine, Institute of Systems Molecular and Integrative Biology (ISMIB), University of Liverpool, Liverpool, United Kingdom

## OPEN ACCESS

### Edited by:

Sandor Kerpel-Fronius,  
Semmelweis University, Hungary

### Reviewed by:

Varvara Barbara Baroutsou,  
Consultant, Athens, Greece  
Kata Mazalin,  
Servier (Hungary), Hungary

### \*Correspondence:

Pierluigi Navarra  
pierluigi.navarra@unicatt.it

### Specialty section:

This article was submitted to  
Drugs Outcomes Research and  
Policies,  
a section of the journal  
Frontiers in Pharmacology

**Received:** 19 January 2022

**Accepted:** 11 March 2022

**Published:** 16 May 2022

### Citation:

Dello Russo C and Navarra P (2022)  
Local Investigators Significantly  
Overestimate Overall Response Rates  
Compared to Blinded Independent  
Central Reviews in Uncontrolled  
Oncology Trials: A Comprehensive  
Review of the Literature.  
Front. Pharmacol. 13:858354.  
doi: 10.3389/fphar.2022.858354

Several drugs gained market authorization based on the demonstration of improved progression-free survival (PFS), adopted as a primary endpoint in Phase 3 clinical trials. In addition, an increasing number of drugs have been granted accelerated approval, and sometimes regular approval, by the main regulatory agencies based on the evaluation of the overall response rate in Phase 1 and 2 clinical trials. However, while the overall survival is an unbiased measure of drug efficacy, these outcomes rely on the assessment of radiological images and patients' categorization using standardized response criteria. The evaluation of these outcomes may be influenced by subjective factors, particularly when the analysis is performed locally. In fact, blinding of treatment is not always possible in modern oncology trials. Therefore, a blinded independent central review is often adopted to overcome the problem of expectation bias associated with local investigator assessments. In this regard, we have recently observed that local investigators tend to overestimate the overall response rate in comparison to central reviewers in Phase 2 clinical trials, whereas we did not find any significant evaluation bias between local investigators and central reviews when considering progression-free survival in both Phase 2 and 3 trials. In the present article, we have tried to understand the reasons behind this discrepancy by reviewing the available evidence in the literature. In addition, a further analysis of Phase 2 and 3 clinical trials that included the evaluation of both endpoints showed that local investigators significantly overestimate overall response rates compared to blinded independent central reviews in uncontrolled oncology trials.

**Keywords:** objective response rate, ORR, progression-free survival, PFS, local investigators, BICR, oncology, pharmacotherapy

## INTRODUCTION

Cancer is one of the major causes of death across the world, leading to a significant reduction of life expectancy in several countries. The burden of cancer is steadily increasing worldwide (Sung et al., 2021); in 2020, there were 19.29 million new cancer cases, 9.96 million cancer deaths, and a total of 50.55 million people living with cancer within 5 years of diagnosis (Global Cancer Observatory, 2020). The COVID-19 pandemic has negatively impacted on several clinical procedures in oncology, such as reduced access to screening programs, delayed diagnosis, and disruption in treatment schedules, possibly contributing to such increase. The rising incidence of cancer coupled with the high mortality rates highlights the need for new and more effective treatments in oncology. However, the outcome of anticancer treatments significantly depends on the extent of tumor disease at treatment initiation, which further highlights the relevance of preventive strategies as well as diagnostic procedures or other clinical interventions to allow early diagnosis of cancer and timely treatment initiation (Kwong et al., 2021; Okoli et al., 2021). With respect to the development of novel anticancer drugs, there is a continuous effort to define the most accurate methods as well as the relevant endpoints to assess the clinical benefit of new compounds in the context of different tumors (Daniele et al., 2020).

In this regard, the overall survival (OS) is an unbiased measure of drug efficacy, and thus it is considered the gold standard endpoint for most types of tumors (Delgado and Guddati, 2021). However, it is not always possible—or convenient—to take the OS as the primary endpoint of drug efficacy in clinical trials (Lebwohl et al., 2009). In fact, several drugs gained market authorization based on the demonstration of improved progression-free survival (PFS), which was adopted as a primary endpoint in Phase 3 clinical trials (Robinson et al., 2014). This is particularly frequent in the context of metastatic disease, although the PFS is not always predictive of beneficial effects in terms of OS (Haslam et al., 2019; Pasalic et al., 2020). Moreover, an increasing number of drugs have been granted accelerated approval, and sometimes regular approval, by the FDA based on the evaluation of the overall response rate (ORR) in Phase 1–2 clinical trials (Chen et al., 2019), although a large meta-analysis showed poor correlation between ORR and OS or PFS (Cooper et al., 2020). In addition, it has become progressively clear, especially with the development of targeted therapies, that not all new drugs cause shrinking of tumors, with disease stabilization observed in several cases. Therefore, it is also often necessary to perform evaluations of time-to-event endpoints during the early phases of clinical development (Gravanis et al., 2014).

Notably, both the ORR and the PFS rely on the assessment of radiological images and patients' categorization based on standardized response criteria (Delgado and Guddati, 2021). The most frequently adopted criteria are the response evaluation criteria in solid tumors (RECIST) (Therasse et al., 2000; Eisenhauer et al., 2009), which have been adequately modified over time to better assess responses in specific clinical settings, for a specific tumor, or in relation to specific

drugs, that is, immune check point inhibitors (Aykan and Özatlı, 2020). A major drawback is that the evaluation of these endpoints may be influenced by subjective factors, such as differences in tumor measurement, differences in the selection of target lesions, failure to diagnose new lesions, and differences in the analysis of non-target lesions (Amit et al., 2010; Iannessi et al., 2021). This may be particularly relevant when the evaluations are performed locally, considering that blinding of treatment is not always possible in modern oncology trials. Therefore, a blinded independent central review (BICR) is often adopted to overcome the problem of expectation bias associated with local investigator (LI) assessments (Dodd et al., 2008; Amit et al., 2010). When BICR is implemented, all the radiological images and selected clinical data acquired in the clinical trials are reviewed by independent radiologists who are blinded to treatment assignment and to any kind of clinical data that can influence the independent review process (Ford et al., 2009). In this regard, we have recently observed that LIs tend to overestimate the ORR in comparison to BICR in Phase 2 clinical trials, whereas we did not find any significant evaluation bias between LIs and BICR when considering PFS in both Phase 2 and 3 trials (Dello Russo et al., 2021; 2020). In the present article, we have tried to understand the reasons behind this discrepancy by reviewing the available evidence in the literature. In addition, we performed a further analysis of Phase 2 and 3 clinical trials that included the evaluation of both endpoints, ORR and PFS.

## LOCAL INVESTIGATOR EVALUATION VERSUS BLINDED INDEPENDENT CENTRAL REVIEW IN THE ASSESSMENT OF RELEVANT PRIMARY OUTCOMES

### Overall Response Rate

The ORR measures the response to a pharmacological treatment through the evaluation of changes in the overall tumor burden in comparison to baseline values (Aykan and Özatlı, 2020). In solid tumors, this is calculated using RECIST criteria (Therasse et al., 2000), updated in 2009 as RECIST v1.1 (Eisenhauer et al., 2009). These criteria allow to classify patients into different groups considering the effects of the investigational drug on the tumor lesions. In brief, patients are categorized as having a complete response (CR), partial response (PR), stable disease (SD), or progressive disease (PD). The ORR is then calculated as the percentage of patients with CR and PR over the total number of eligible patients, that is, all subjects included in the trial (Eisenhauer et al., 2009). Response evaluation is usually based on arbitrary cutoff values at a specific time point, usually at 4 or 8 weeks after treatment initiation (Eisenhauer et al., 2009; Aykan and Özatlı, 2020).

In the first analysis of 18 Phase 3 clinical trials, a significant variability was reported between the LI assessments, and the BICR results in the evaluation of the tumor response rate (RR) (Tang et al., 2010). Despite using a general abbreviation, in the majority of the analyzed trials, the RR is indeed equal to the ORR,

**TABLE 1** | Main characteristics of Phase 2 clinical trials selected for the analysis.

	<b>Trial registration number (First posted date)</b>	<b>Intervention model/ masking</b>	<b>Tumor type</b>	<b>Treatment arm(s)</b>	<b>Number of patients</b>	<b>Primary endpoint</b>	<b>Cycle length</b>	<b>Timing of scans</b>	<b>Time frame (primary endpoint)</b>	<b>DI ORR</b>	<b>DI PFS</b>	<b>Reference</b>
1	NCT00875667 (3 Apr 2009)	Parallel assignment/ none (open label)	Relapsed or refractory mantle cell lymphoma	E: lenalidomide AC: investigator's choice single agent	E: 170 AC: 84	PFS per BICR and LI <sup>a</sup>	E: 28 days AC: variable according to different regimens	Every 2 cycles for 6 months; every 90 days thereafter until disease progression	Median follow-up of 15.9 months (IQR 731.7) for all patientsa	E: 1.15 AC: 2.09	E: 0.99 AC: 1.04	Trněný et al. (2016)
2	NCT01685060 (13 Sep 2012) EudraCT 2012-003432-24 (12 Apr 2017)	Single Group Assignment/ None (Open Label)	Non-small-cell lung cancer	Ceritinib (LDK378)	140	ORR per LI	28 days	Every 8 weeks	Up to 24 weeks	1.14	0.78	Hida et al. (2017)
3	NCT01685138 (14 Sep 2012)	Single group assignment/ none (open label)	Non-small-cell lung cancer	Ceritinib (LDK378)	124	ORR per LI	28 days	Every 8 weeks	Up to 5 years	1.06	0.86	Nishio et al. (2020)
4	NCT01708174 (16 Oct 2012)	Single group assignment/ none (open label)	Medulloblastoma	Sonidegib (LDE225) and TMZ	16	ORR per BICR	28 days	NA	>3 years	1.33	1	Clinicaltrials.gov identifier NCT01708174, 2012
5	NCT01984242 (14 Nov 2013) EudraCT 2013-003167-58 (29 Oct 2017)	Parallel assignment/ none (open label)	Renal cell carcinoma	ITT E1: atezolizumab + bevacizumab E2: atezolizumab AC: sunitinib	ITT E1: 101 E2: 103 AC: 101	PFS per BICR	6 weeks E(1-2): one infusion every 3 weeks AC: Day 1-28 of each cycle	Every 12 weeks ± 5 days after cycle 1	Up to ~2.75 years	ITT E1: 1.09 E2: 0.92 AC: 1.14	ITT E1: 0.95 E2: 0.90 AC: 0.93	McDermott et al. (2018)
6	NCT02031458 (9 Jan 2014) EudraCT 2013-003330-32 (1 July 2016)	Single group assignment/ none (open label)	Non-small-cell lung cancer	Atezolizumab Cohort 1: first line Cohort 2: second line Cohort 3: third line and beyond	Cohort 1: 139 Cohort 2: 267 Cohort 3: 253	ORR per BICR	3 weeks	Every 6 weeks for 12 months; every 9 weeks thereafter	Up to 16 months	Cohort 1: 1.15 Cohort 2: 1.09 Cohort 3: 1.05	Cohort 1: 1.29 Cohort 2: 1.00 Cohort 3: 1.07	Peters et al. (2017)
7	NCT02040870 (20 Jan 2014)	Single group assignment/ none (open label)	Non-small-cell lung cancer	Ceritinib (LDK378)	103	Others	28 days	Every 8 weeks	40 months	1.30	1.89	Wu et al. (2020)
8	NCT02336451 (13 Jan 2015) EudraCT 2014-000578-20 (22 Feb 2020)	Parallel assignment/ none (open label)	ALK-positive Non-small-cell lung cancer	Ceritinib <sup>c</sup> (LDK378)	Arm 1: 42 Arm 2: 40 Arm 3: 12 Arm 4: 44 Arm 5: 18	ORR per LI	28 days	At week 8; every 8 weeks thereafter	43 months	Arm 1: 1.50 Arm 2: 2.00 Arm 3: 1.50 <sup>b</sup> Arm 4: 0.96 Arm 5: 1.50	Arm 1: 1.44 Arm 2: 1.02 Arm 3: NA <sup>b</sup> Arm 4: 1.03 Arm 5: 1.44	Chow et al. (2019)

(Continued on following page)

**TABLE 1 |** (Continued) Main characteristics of Phase 2 clinical trials selected for the analysis.

	<b>Trial registration number (First posted date)</b>	<b>Intervention model/ masking</b>	<b>Tumor type</b>	<b>Treatment arm(s)</b>	<b>Number of patients</b>	<b>Primary endpoint</b>	<b>Cycle length</b>	<b>Timing of scans</b>	<b>Time frame (primary endpoint)</b>	<b>DI ORR</b>	<b>DI PFS</b>	<b>Reference</b>
9	NCT00117637 (8 July 2005)	Parallel assignment/ none (open label)	Renal cell carcinoma	E: first sorafenib (BAY43-9006) 400 mg then 600 mg AC: first interferon then sorafenib 400 mg	E: 97 AC: 92	PFS per BICR	NA	Every 8 weeks	~15 months	E: 4.19 AC: 1.75	E: 0.98 AC: 1.25	Escudier et al. (2009)
10	NCT01943864 (September 17, 2013)	Single group assignment/ none (open label)	Advanced or metastatic biliary tract cancers in Japanese	Trametinib	20	Others	NA	Every 8 weeks	Up to week 37	NA <sup>b</sup> (ORR per LI: 0)	1 <sup>b</sup>	Ikeda et al. (2018)
11	NCT00679211 (16 May 2008)	Single group assignment/ none (Open Label)	Metastatic breast cancer	Trastuzumab emtansine (Kadcyla)	110	ORR per BICR	3 weeks	Every other cycle by LI and retrospectively by double-reader BICR as needed	~9 months	1	0.797	Krop et al. (2012)
12	NCT02674061 (4 Feb 2016)	Parallel assignment/ none (open label)	Advanced recurrent ovarian cancer	Pembrolizumab (MK-3475) Cohort A (0–2 prior lines of treatment) Cohort B (3–5 prior lines of treatment)	Cohort A: 285 Cohort B: 91	ORR per BICR (in all cohorts A and B and subgroup cohorts PDL-L1+)	3 weeks	Every 9 weeks for the first 54 weeks; every 12 weeks thereafter	Up to ~43 months	Cohort A: 0.864 Cohort B: 0.889	Cohort A: 1 Cohort B: 1	Matulonis et al. (2019)
13	NCT03525678 (16 May 2018) EudraCT 2017-004810-25 (26 Apr 2020)	Parallel assignment/ none (open label)	Multiple myeloma	Belantamab mafodotin (GSK2857916) Arm 1 (2.5 mg/kg frozen liquid, FL) Arm 2 (3.4 mg/kg FL) Arm 3 (3.4 mg/kg lyophilized powder, LP)	Arm 1: 97 Arm 2: 99 Arm 3: 25	ORR per BICR	3 weeks	Every 3 weeks (radiography for skeletal lytic lesions, but it is not specified in the timing)	Up to 48 weeks	Arm 1: 0.968 Arm 2: 0.912 Arm 3: 1.083 <sup>b</sup>	Arm 1: 0.759 Arm 2: 0.776 Arm 3: NA <sup>b</sup>	Lonial et al. (2020)
14	NCT02576990 (15 Oct 2015)	Single group assignment/ none (open label)	Mediastinal large B-cell lymphoma (rrMLBCL) Richter syndrome (rrRS)	Pembrolizumab (MK-3475) Arm 1 (rrMLBCL) Arm 2 (rrRS)	Arm 1: 53 Arm 2: 23	ORR per BICR	3 weeks	At week 12 and then every 12 weeks	Up to ~27 months	Arm 1: 0.916 Arm 2: 0.331	Arm 1: 0.782 Arm 2: 1.125	Armand et al., 2019; Armand et al., 2020

(Continued on following page)

**TABLE 1 |** (Continued) Main characteristics of Phase 2 clinical trials selected for the analysis.

	<b>Trial registration number (First posted date)</b>	<b>Intervention model/ masking</b>	<b>Tumor type</b>	<b>Treatment arm(s)</b>	<b>Number of patients</b>	<b>Primary endpoint</b>	<b>Cycle length</b>	<b>Timing of scans</b>	<b>Time frame (primary endpoint)</b>	<b>DI ORR</b>	<b>DI PFS</b>	<b>Reference</b>
15	NCT01660451 (8 Aug 2012)	Parallel assignment/ none (open label)	Non-Hodgkin's lymphoma (NHL)	Copanlisib (BAY80-6946) Part A (indolent NHL/CLL) Part A (aggressive NHL) Part B (indolent B-cell NHL)	Part A (indolent): 33 Part A (aggressive): 51 Part B: 142	ORR per BICR	28 days	Every two cycles during year 1; every three cycles during year 2; every six cycles during year 3	Up to 16 weeks of treatment	Part A (indolent): 1.072 Part A (aggressive): 1.154 Part B: 0.869	Part A (indolent): 0.952 Part A (aggressive): 1 Part B: 0.971	Dreyling et al. (2017) Panayiotidis et al. (2021)
16	EudraCT 2013-005486-39 (28 July 2016) NCT02108652 (9 Apr 2014) NCT02951767 (1 Nov 2016)	Single group assignment/ none (open label)	Urothelial bladder cancer	Atezolizumab Cohort 2: second-line or beyond treatment Cohort 1: drug naive	Cohort 2: 310 Cohort 1: 119	ORR per LI and BICR	21 days	Every 9 weeks for the first 12 months; every 12 weeks thereafter	Up to maximum length of follow-up of 24.5 months	1.04 1.11	1 1.55	Perez-Gracia et al. (2018) Balar et al. (2017)
17	EudraCT 2013-002737-38 (28 Apr 2021) NCT02183870 (no results) (8 Jul 2014)	Single group assignment/ none (open label)	Lung cancer Adenocarcinoma NSCLC (ROS1 +)	Crizotinib	Per protocol: 30 ITT: 34	ORR per LI	Continuous daily dosing	Every 6 weeks	NA	Per protocol: 0.959 ITT: 0.959 <sup>b</sup>	Per protocol: 0.97 ITT: NA <sup>b</sup>	Michels et al. (2019)

*Trials that were included in our initial analysis (Dello Russo et al., 2021) are highlighted in gray. IQR, interquartile range; ITT, intention to treat population; ORR, objective response rate; PFS, progression-free survival; BICR, blind independent central review; LI, local investigator; rMLBCL, relapsed or refractory primary mediastinal large B-cell lymphoma; rRS, relapsed or refractory richter syndrome; DI, discrepancy index; NHL, non-Hodgkin's lymphoma; CLL, chronic lymphocytic leukemia; mRECIST, modified response evaluation criteria in solid tumors; PFI/TFI, platinum-free interval/treatment-free interval; PrALKi, previous ALK inhibitors' treatment; PrBRad, previous brain radiations; RECIST, response evaluation criteria in solid tumors.*

<sup>a</sup>The Primary outcome for this trial was PFS per BICR (up to data cutoff date, 7 March 2014) and LI (at final analysis, up to study discontinuation of 09 October 2018, (median follow up of 285 weeks). The data included in our analysis refer to the primary analysis cutoff date for both LI and BICR (Trněný et al., 2016). The same DIs are calculated by comparing the values obtained by LI at final analysis versus BICR primary analysis (Dello Russo et al., 2021).

<sup>b</sup>Not included in the present analysis since the specific arm/group was missing the DI for one of the outcomes (either ORR or PFS).

<sup>c</sup>Treatment arms were the followings: Arm 1 (PrALKi = Yes; PrBRad=Yes); Arm 2 (PrALKi=Yes; PrBRad=No); Arm 3 (PrALKi=No; PrBRad=Yes); Arm 4 (PrALKi=No; PrBRad=No); Arm 5 (leptomeningeal carcinomatosis). In Arm 1–4, no evidence of leptomeningeal carcinomatosis.

**TABLE 2 |** Comparison between time to first response (TTFR), time to disease progression (TTP), and median PFS time in Phase 2 clinical trials.

Trial registration number	Treatment arm(s)	Timing of scans	Number of patients (TTFR)	TTFR per LI (months, 95% CI)	TTFR per BICR (months, 95% CI)	Number of patients (PFS)	Time to progression (TTP) per LI (months, 95% CI)	Time to progression (TTP) per BICR (months, 95% CI)	Median PFS per LI (months, 95% CI)	Median PFS per BICR (months, 95% CI)	
1	NCT00875667	E: lenalidomide AC: investigator's choice single agent	Every two cycles for 6 months; every 90 days thereafter until disease progression <sup>a</sup>	E: 170 AC: 84	E: 5.5 (3.9 – 5.9) AC: 9.2 (5.9 – NR)	E: 4.3 (3.9 – 11.5) AC: NR	E: 170 AC: 84	E: 9.1 (5.8 – 14.1) AC: 5.7 (3.7 – 8.5)	E: 9.1 (6.6 – 12.2) AC: 5.7 (3.7 – 6.9)	E: 8.7 (5.5 – 12.1) AC: 5.2 (3.7 – 6.9)	
2	NCT01685060 - EudraCT 2012-003432-24	Ceritinib (LDK378)	Every 8 weeks	LI: 57 BICR: 50	3.0 (SD, 3.54)	2.2 (SD, 1.44)	140	NA	NA	5.8 (5.4–7.6)	7.4 (5.6–10.9)
3	NCT01685138	Ceritinib (LDK378)	Every 8 weeks	LI: 84 BICR: 79	2.5 (SD, 2.66)	2.2 (SD, 1.22)	124	NA	NA	16.6 (11.0–23.2)	19.4 (10.9–29.3)
7	NCT02040870	Ceritinib (LDK378)	Every 8 weeks	LI: 43 BICR: 33	1.90 (1.6–12.9)	1.80 (1.6–3.7)	103	NA	NA	7.2 (4.1–7.5)	3.8 (3.6–5.6)
8	NCT02336451	Ceritinib (LDK378)	At week 8; every 8 weeks thereafter	<b>LI</b> Arm 1: 15 Arm 2: 12 Arm 3: 6 Arm 4: 26 Arm 5: 3 <b>BICR</b> Arm 1: 10 Arm 2: 6 Arm 3: 4 Arm 4: 27 Arm 5: 2	Arm 1: 1.87 (1.7–9.3) Arm 2: 2.00 (1.7–9.3) Arm 3: 1.82 (1.2–30.1) Arm 4: 1.81 (1.3–3.7) Arm 5: 1.91 (1.8–3.6)	Arm 1: 2.00 (1.7 – 12.9) Arm 2: 1.76 (1.6–1.9) Arm 3: 1.82 (1.7–26.5) Arm 4: 1.81 (1.3–22.0) Arm 5: 1.86 (1.8–1.9)	<b>LI</b> Arm 1: 32 Arm 2: 35 Arm 3: 6 Arm 4: 33 Arm 5: 14 <b>BICR</b> Arm 1: 34 Arm 2: 36 Arm 3: 8 Arm 4: 33 Arm 5: 14	NA	NA	Arm 1: 7.2 (3.3–10.9) Arm 2: 5.6 (3.6–9.2) Arm 3: NR (1.0 - NR) Arm 4: 7.9 (5.5–9.4) Arm 5: 5.2 (1.6–7.2)	Arm 1: 5.0 (3.3–9.1) Arm 2: 5.5 (3.6–7.3) Arm 3: 15.5 (1.0 - NR) Arm 4: 7.7 (5.5–9.7) Arm 5: 3.6 (1.6–5.4)
9	NCT00117637	Sorafenib (BAY43-9006) + interferon	Every 8 weeks	<b>LI</b> E: 21 AC: 14 <b>BICR</b> E: 5 AC: 8	E: 3.5 (1.6–11.1) AC: 5.4 (1.2–18.3)	E: 1.8 (1.7–3.7) AC: 5.4 (3.7–11)	E: 97 AC: 92	NA	NA	E: 5.6 (5.4–7.5) AC: 7.0 (5.4–8.8)	E: 5.7 (5.0–7.4) AC: 5.6 (3.7–7.4)
10	NCT01943864	Trametinib	Every 8 weeks	LI: 0 BICR: 1	NA	20.1 weeks	20	NA	NA	10.6 weeks (4.6–12.1)	10.6 weeks (4.6–12.7)
13	NCT03525678 (EudraCT 2017-004810-25)	Belantamab mafodotin (GSK2857916)	Every 3 weeks (radiography for skeletal lytic lesions, but it is not specified in the timing)	<b>LI</b> Arm 1: 29 Arm 2: 31 Arm 3: 13 <b>BICR</b> Arm 1: 30 Arm 2: 34 Arm 3: 12	Arm 1: 1.4 (0.8–2.1) Arm 2: 1.5 (0.9–3.0) Arm 3: 0.9 (0.8–1.0)	Arm 1: 1.4 (0.8–2.1) Arm 2: 1.4 (0.8–2.8) Arm 3: 0.9 (0.8–1.6)	Arm 1: 97 Arm 2: 99 Arm 3: 25	Arm 1: 2.3 (0.8 - NR) Arm 2: 4.2 (1.3 - NR) Arm 3: 4.3 (2.1 - NR)	Arm 1: 3.0 (0.9 - NR) Arm 2: 5.8 (0.9 - NA) Arm 3: NA (2.2 - NA)	Arm 1: 2.2 (0.8 - NR) Arm 2: 3.8 (1.1 - NR) Arm 3: 4.3 (2.1 - NR)	Arm 1: 2.9 (0.9 - NR) Arm 2: 4.9 (0.9 - NR) Arm 3: NA (2.2 - NR)

*Trials that were included in our initial analysis (Dello Russo et al., 2021) are highlighted in gray. ITT, intention to treat population; NA, not available; NR, not reached; ORR, objective response rate; PFS, progression-free survival; BICR, blind independent central review.*

<sup>a</sup>Data included in the table refer to the analysis at data cutoff date, 7 March 2014 (Trněný et al., 2016).

**TABLE 3 |** Main characteristics of Phase 3 clinical trials selected for the analysis.

	<b>Trial registration number (first posted date)</b>	<b>Intervention model/masking</b>	<b>Tumor type</b>	<b>Treatment arm(s)</b>	<b>Number of patients</b>	<b>Primary endpoint</b>	<b>Cycle length</b>	<b>Timing of scans</b>	<b>Time frame (primary endpoint)</b>	<b>DI ORR</b>	<b>DI PFS</b>	<b>Reference</b>
1	NCT00069108 (September 18, 2003)	Parallel assignment/ none (open label)	Colorectal cancer	E: XELOX AC: FOLFOX-4	E (ITT): 313 AC (ITT): 314 E (PP): 251 AC (PP): 314	PFS per LI	E: 3 weeks up to eight cycles AC: 2 weeks up to 12 cycles	Every 6 weeks (+ within 2 weeks of study completion, withdrawal or treatment discontinuation)	Up to 3 years	E (ITT): 1.50 AC (ITT): 1.33	E (PP): 0.92 AC (PP): 1.04	Rothenberg et al. (2008)
2	NCT02370498 (25 Feb 2015)	Parallel assignment/ none (open label)	Gastric adenocarcinoma Gastroesophageal junction adenocarcinoma	E: pembrolizumab AC: paclitaxel	E (all): 296 AC (all): 296 E (PD-L1+): 196 AC (PD-L1+): 199	PFS (and OS) per BICR in PD-L1+ patients	E: 21 days AC: 28 days	Every 6 weeks	Up to 30 months	E (all): 1.10 AC (all): 1.22 E (PD-L1+): 1.09 <sup>a</sup> AC (PD-L1+): 1.15 <sup>a</sup>	E (all): 1.07 AC (all): 0.78 E (PD-L1+): 1.07 <sup>a</sup> AC (PD-L1+): 0.76 <sup>a</sup>	Shitara et al. (2018)
3	NCT00075270 (9 Jan 2004)	Parallel assignment/ double (participant, investigator)	Metastatic breast cancer	E: lapatinib (+ paclitaxel) AC: placebo (+ paclitaxel)	E: 291 AC: 288	TTP per LI and BICR	E: 3 weeks AC: 3 weeks	For efficacy 9 weeks after study entry, at 12-week intervals, and at treatment end. For survival at 12-week intervals	Average 26 weeks	E: 1.31 AC: 1.35	E: 0.86 <sup>b</sup> AC: 0.88 <sup>b</sup>	Di Leo et al. (2008)
4	NCT01120184 (10 May 2010)	Parallel assignment/ triple (participant, investigator, and outcomes assessor)	Breast cancer	AC: trastuzumab + taxane E1: trastuzumab emtansine + placebo E2: trastuzumab emtansine + pertuzumab	PFS AC: 365 E1: 367 E2: 363 ORR AC: 287 E1: 303 E2: 299	PFS per BICR	3 weeks (except paclitaxel every 1 week)	Every 9 weeks for 81 weeks, then every 12 weeks thereafter, and/or up to 42 days after last dose	Up to 48 months	AC: 1.02 E1: 1.08 E2: 1.05	AC: 0.91 E1: 1 E2: 0.97	Perez et al. (2017)
5	NCT00689936 (4 Jun 2008)	Parallel assignment/ none (open label)	Multiple myeloma (previously untreated; stem cell transplant ineligible)	Arm 1: lenalidomide + low-dose DEX (until disease progression) Arm 2: lenalidomide + low-dose DEX (18 cycles) Arm 3/AC: melphalan + prednisone + thalidomide	Arm 1: 535 Arm 2: 541 Arm 3/AC: 547	PFS per BICR and LI	4 weeks	After each treatment cycle and every 28 days during the follow-up phase	PFS by BICR: median follow-up time of 17.1 months PFS by BICR: median follow-up time of 17.7 months	Arm 1: 1.07 Arm 2: 1.07 Arm 3/AC: 1.08	Arm 1: 1.02 Arm 2: 1.01 Arm 3/AC: 1.03	Benboubker et al. (2014)

(Continued on following page)

**TABLE 3 |** (Continued) Main characteristics of Phase 3 clinical trials selected for the analysis.

	<b>Trial registration number (first posted date)</b>	<b>Intervention model/ masking</b>	<b>Tumor type</b>	<b>Treatment arm(s)</b>	<b>Number of patients</b>	<b>Primary endpoint</b>	<b>Cycle length</b>	<b>Timing of scans</b>	<b>Time frame (primary endpoint)</b>	<b>DI ORR</b>	<b>DI PFS</b>	<b>Reference</b>
6	NCT01360554 (25 May 2011)	Parallel assignment/ quadruple (participant, care provider, investigator, and outcomes assessor)	Non-small-cell lung cancer	E: dacomitinib (PF-00299804) + placebo (erlotinib) AC: erlotinib + placebo (PF- 00299804)	All population E: 439 AC: 439	PFS per BICR and PFS in KRAS wild- type patients	28 days (continuous oral daily dosing)	At the end of cycles 2, 3, and 4, then every other cycle	Median follow- up of 7.1 months, until disease progression	E: 1.13 AC: 1.29	E: 0.73 AC: 0.76	Ramalingam et al. (2014)
7	NCT01774721 (24 Jan 2013)/ EudraCT 2012-004977- 23 (25 Oct 2018)	Parallel assignment/ none (open label)	Non-small-cell lung cancer with EGFR- activating mutations	E: dacomitinib (PF-00299804) AC: gefitinib	E: 227 AC: 225	PFS per BICR	28 days (continuous oral daily dosing)	At the end of cycles 1–2, then at every other cycle	Up to 48 months	E: 1.01 AC: 0.98	E: 1.13 AC: 1.20	Wu et al. (2017)
8	NCT02604342 (13 Nov 2015)/ EudraCT 2015- 000634-29	Parallel assignment/ none (open label)	Non-small-cell lung cancer	E: alectinib AC: premetrexed/ docetaxel	E: 79 AC: 40	PFS per LI	3 weeks (alectinib: continuous oral twice daily dosing)	Every 6 weeks	Up to 33 months	E: 1.40 AC: 0.22	E: 1.35 AC: 0.875	Novello et al. (2018)
9	NCT01245062 (22 Nov 2010)	Crossover assignment/ none (open label)	Melanoma	E: trametinib (GSK1120212) AC: dacarbazine or paclitaxel	BRAF V600E + w/o brain metastasis E: 178 AC: 75	PFS in BRAF V600E+ w/o brain metastasis per BICR and LI	3 weeks (trametinib: continuous dosing)	At weeks 6, 12 21, and 30; then, every 12 weeks	Average of 20.3 months	BRAF V600E+ w/o brain metastasis E: 1.30 AC: 2.33	BRAF V600E+ w/o brain metastasis E: 0.92 AC: 0.88	Flaherty et al. (2012)
10	NCT02718417 (24 Mar 2016) EudraCT 2015- 003239-36	Parallel assignment/ none (open label)	Ovarian cancer	AC: chemotherapy then observation E1: chemotherapy then avelumab in maintenance E2: chemotherapy in combination with avelumab then avelumab in maintenance	AC: 335 E1: 332 E2: 331	PFS per BICR	Chemotherapy: 3 weeks Avelumab: 2 weeks	After three cycles and at completion of chemotherapy; then, every 12 weeks during maintenance	Maximum duration of 27 months	AC: 0.914 <sup>a</sup> E1: 0.852 E2: 0.864	AC: NA <sup>a</sup> E1: 0.821 E2: 0.890	Monk et al. (2021)
11	NCT00083889 (4 Jun 2004)	Parallel assignment/ none (open label)	Renal cell carcinoma	AC: IFN $\alpha$ E: sunitinib (SU011248)	AC: 375 E: 375	PFS per BICR and LI	AC: 3 weeks E: 6 weeks	At day 28 of cycles 1 through 4, and every two cycles thereafter until the end of treatment	Duration of treatment phase	AC: 1.50 E: 1.19	AC: 1.00 E: 0.99	Motzer et al. (2007)

(Continued on following page)

**TABLE 3 |** (Continued) Main characteristics of Phase 3 clinical trials selected for the analysis.

	<b>Trial registration number (first posted date)</b>	<b>Intervention model/ masking</b>	<b>Tumor type</b>	<b>Treatment arm(s)</b>	<b>Number of patients</b>	<b>Primary endpoint</b>	<b>Cycle length</b>	<b>Timing of scans</b>	<b>Time frame (primary endpoint)</b>	<b>DI ORR</b>	<b>DI PFS</b>	<b>Reference</b>
12	NCT02421588 (April 20, 2015) EudraCT 2014-005251-39 (17 Oct 2019)	Parallel assignment/ none (open label)	Ovarian cancer (platinum resistant)	E (Arm A): lurbinectedin (PM01183) AC (Arm B): pegylated liposomal doxorubicin or topotecan	E: 221 AC: 221	PFS per BICR	E: 3 weeks AC: pegylated liposomal doxorubicin (4 weeks) topotecan (3 weeks)	Every 8 weeks	Up to 3 years	E: 1.09 AC: 1.31	E: 1.10 AC: 1.00	Gaillard et al. (2021)
13	NCT03052608 (14 Feb 2017)	Parallel assignment/ none (open label)	Non-small-cell lung cancer	E: lorlatinib AC: crizotinib	E: 149 AC: 147	PFS per BICR	28 days	Every 8 weeks (±1 week)	Up to 33 months	E: 1.06 <sup>a</sup> AC: 1.07	E: NA <sup>a</sup> AC: 0.98	Shaw et al. (2020)
14	NCT01102426 (13 Apr 2010)	Parallel assignment/ none (open label)	Relapsed/refractory multiple myeloma	E: plitidepsin + dexamethasone AC: dexamethasone	E: 171 AC: 84	PFS per BICR	4 weeks	NA	Up to 5 years	E: 1.31 AC: 0.33	E: 1.12 AC: 0.65	Spicka et al. (2019)
15	NCT01287741 <sup>c</sup> (1 Feb 2011) EudraCT 2010-024194-39 (23 Apr 2017)	Parallel assignment/ none (open label)	Diffuse large B-cell lymphoma	E: obinutuzumab + chemotherapy AC: rituximab + chemotherapy	E: 712 AC: 706	PFS per LI	21 days	4–8 weeks (CT) or 6–8 weeks (FDG-PET) after the last treatment or sooner in the case of early discontinuation	LE up to approximately 6.5 years BICR: up to approximately 4 years and 9 months	E: 0.999 <sup>a</sup> AC: 0.992 <sup>a</sup>	NA <sup>a</sup>	Vitolo et al. (2017)
16	NCT02580058 <sup>d</sup> (20 Oct 2015)	Parallel assignment/ none (open label)	Ovarian cancer	E1: avelumab E2: avelumab plus pegylated liposomal doxorubicin (PLD) AC: PLD alone	E1: 188 E2: 188 AC: 190	PFS per BICR and OS	Avelumab: 2 weeks; doxorubicin: 4 weeks	MRI or CT scans every 8 weeks	Up to 30 months	E1: 1.43 E2: 1.40 AC: 2.26	E1: 1.00 E2: 1.27 AC: 1.06	Pujade-Lauraine et al. (2021)
17	NCT02603432 <sup>d</sup> (11 Nov 2015)	parallel assignment/ none (open label)	Urothelial cancer	E (Arm A): avelumab plus best supportive care (BSC) Arm B: best supportive care (BSC) alone	E (Arm A): 350 Arm B: 350	OS	4 weeks	Every 8 weeks for 12 months and then every 12 weeks	Up to 41 months at the time of final analysis	E (Arm A): 1.27 Arm B: 2.43	E (Arm A): 1.5 Arm B: 1.05	Powles et al. (2020)
18	EudraCT 2014-004684-20 NCT02420821 (20 Apr 2015)	Parallel assignment/ none (open label)	Untreated advanced renal cell carcinoma	AC: sunitinib E: atezolizumab + bevacizumab	ITT: AC: 461 E: 454	PFS per LI in PD-L1 selected population	AC: 4 weeks on, 2 weeks off E: 3 weeks	At week 12, then every 6 weeks up to week 78, and then every 12 weeks	Up to approximately 24 months	ITT AC: 1.064 E: 1.099	ITT AC: 1.012 E: 1.167	Rini et al. (2019)

(Continued on following page)

**TABLE 3 |** (Continued) Main characteristics of Phase 3 clinical trials selected for the analysis.

	Trial registration number (first posted date)	Intervention model/masking	Tumor type	Treatment arm(s)	Number of patients	Primary endpoint	Cycle length	Timing of scans	Time frame (primary endpoint)	DI ORR	DI PFS	Reference
19	EudraCT 2010-024132-41 (16 Mar 2017) NCT01332968 (11 Apr 2011)	Parallel assignment/ none (open label)	Non-Hodgkin's lymphoma	E: obinutuzumab + chemotherapy AC: rituximab + chemotherapy	Follicular lymphoma population E: 601 AC: 601	PFS per LI in the follicular lymphoma population	21 or 28 days	After three cycles (bendamustine treated) or four cycles (CHOP or CVP) and on the completion of induction therapy; every 2 months for 2 years; then, every 3–6 months, with CT performed every 6–12 months, until progression or withdrawal from the trial	Up to ~4 years and 7 months	E: 0.957 AC: 0.970	E: 1.15 AC: 1.08	Hiddemann et al. (2018); Marcus et al. (2017)
20	EudraCT 2016-004387-18 <sup>d</sup> (01 May 2020) NCT03273153 (6 Sep 2017)	Parallel assignment/ none (open label)	Advanced BRAFV600 wild-type melanoma	E: cobimetinib + atezolizumab AC: pembrolizumab	E: 222 AC: 224	PFS per BICR	Every 8 weeks through 80 weeks; then, every 12 weeks until progression	3 weeks	For approximately 16 months	E: 1.07 AC: 1.16	E: 1.02 AC: 1.26	Gogas et al. (2021)
21	NCT00789373 (First posted: 11 Nov 2008)	Parallel assignment/ quadruple (participant, care provider, investigator, and outcomes assessor)	Non-small-cell lung cancer	E: pemetrexed (maintenance) AC: placebo	E: 316/359 AC: 156/180	PFS per LI	Every other cycle (6 weeks [±1])	21 days	(Up to 19.3 months)	E: 1 AC: 1	E: 1.043 AC: 1.088	Paz-Ares et al. (2012)

Trials that were included in our initial analysis (Dello Russo et al., 2021) are highlighted in gray. ITT, intention to treat population; ORR, objective response rate; PFS, progression-free survival; PD-L1, programmed cell death ligand 1; PP, per protocol; TTP, time to progression.

<sup>a</sup>Not included in the analysis because the subgroup of the whole population or the specific arm/group was missing the DI for one of the outcomes (either ORR or PFS).

<sup>b</sup>Time to progression was used for comparative analysis.

<sup>c</sup>Trial NCT01287741 is included in the table because HR values for PFS per LI and BICR were available for comparison. DI was calculated and included in a pooled analysis (see text).

<sup>d</sup>For these trials, DI based of the HR values for PFS was not available.

**TABLE 4 |** Comparison between time to first response (TFR), time to disease progression (TTP), and median PFS in Phase 3 clinical trials.

Trial registration number	Treatment arm(s)	Timing of scans	Number of patients (TTFR)	TTFR per LI (months, 95% CI)	TTFR per BICR (months, 95% CI)	Number of patients	Time to progression (TTP) per LI (months, 95% CI)	Time to progression (TTP) per BICR (months, 95% CI)	Median PFS per LI (months, 95% CI)	Median PFS per BICR (months, 95% CI)
5	NCT00689936	Arm 1: lenalidomide + low-dose DEX (until disease progression) Arm 2: lenalidomide + low-dose DEX (18 cycles) Arm 3/AC: melphalan + prednisone + thalidomide	After each treatment cycle and every 28 days during the follow-up phase	Arm 1: 402 Arm 2: 397 Arm 3/AC: 341	Arm 1: 1.8 (0.50–22.2) Arm 2: 1.8 (0.8–34.8) Arm 3/AC: 2.8 (1.2–56.3)	Arm 1: 535 Arm 2: 541 Arm 3/AC: 547	NA	NA	Arm 1: 26.0 (20.7–29.7) Arm 2: 21.0 (19.7–22.4) Arm 3/AC: 21.9 (19.8–23.9)	Arm 1: 25.5 (20.7–29.4) Arm 2: 20.7 (19.4–22.0) Arm 3/AC: 21.2 (19.3–23.2)
13	NCT03052608	E: lorlatinib AC: crizotinib	Every 8 weeks (±1 week)	E: 113 AC: 85	NA	E: 149 AC: 147	NA	NA	E: NR (NR to NR) AC: 9.1 (7.4–10.9)	E: NR (NR to NR) AC: 9.3 (7.6–11.1)
17	NCT02603432	E (Arm A): avelumab plus best supportive care (BSC) Arm B: best supportive care (BSC) alone	Every 8 weeks for 12 months and then every 12 weeks	LI E (Arm A): 43 Arm B: 12 BICR E (Arm A): 34 Arm B: 5	E (Arm A): 2.0 (1.8–22.2) Arm B: 1.9 (1.1–10.9)	E (Arm A): 2.0 (1.7–16.4) Arm B: 2.0 (1.8–7.0)	NA	NA	E (Arm A): 5.5 (4.2–7.2) Arm B: 2.1 (1.9–3.0)	E (Arm A): 3.7 (3.5–5.5) Arm B: 2.0 (1.9–2.7)

*Trials that were included in our initial analysis (Dello Russo et al., 2021) are highlighted in gray. ITT, intention to treat population; ORR, objective response rate; PFS, progression-free survival; BICR, blind independent central review; NA, not available; NR, not reached.*

per the abovementioned definition. However, in two trials, CR was observed only in one enrolled patient, which leaves the comparison among LI and BICR to the evaluation of the rate of PRs (Escudier et al., 2007; Motzer et al., 2007). In 14 out of 18 trials, the RR values measured by LIs were consistently greater than those by BICR for both the experimental and the control arms of the studies. For the remaining four trials, in one study, the evaluation of the RR was higher in the experimental group by LIs than that by BICR and lower for the control arm; in two trials, the RR was reported to be inferior by LIs in comparison to BICR in the experimental group and superior in the Control arm; and in one trial, the RR was consistently lower in both arms of the study by the LI in comparison to BICR. The analysis was performed by comparing the absolute values reported for the RR. Therefore, the RR reported by LIs was greater in 31 out of 36 evaluations, with an estimated mean difference of +4.57% (95% confidence interval, CI, 2.95–6.19%), thus suggesting that overall the LIs were more “optimistic” than the independent reviewers in the evaluation of tumor RR. These results are consistent with a more recent analysis carried out by our research group on data from 20 Phase 2 clinical trials having the ORR as primary or secondary endpoint, assessed by both LIs and through BICR (Dello Russo

et al., 2021). In some trials, more than one treatment group was included for a total of 33 ORR values. In most cases (26/33, 78.8%), the LI assessment was more “optimistic” in the evaluation of ORR, with an average of + 17.5% positive responses than the central independent reviewer. Most trials had an open-label design, and 12 out of 20 trials (60%) had a single group assignment. When parallel groups were included, these referred, for example, to different doses of the same drugs (NCT02094573, brigatinib 90 mg, and brigatinib 90–180 mg) (Kim et al., 2017) or to the evaluation of the same drug in different cohorts of patients (NCT02336451, ceritinib in different cohorts of patients with brain metastases, no evidence of leptomeningeal carcinomatosis, and exposed to different previous therapies) (Chow et al., 2019). Therefore, most trials lacked a comparator arm. Our data are however consistent with previous evidence from the evaluation of 28 Phase 3 clinical trials, six of which were carried out according to a blinded design (Zhang et al., 2017). In this study, with respect to the analysis of ORR on both the experimental and control arms of the studies, the authors found a lower response frequency in the central review compared with LI assessment. The discrepancy was larger in the open-label trials, for a positive primary outcome, and

central assessed primary endpoint. However, this phenomenon impacted both the experimental and the comparator arm; therefore, it did not cause any significant evaluation bias between the two reviews. The ratio of odd ratios (ORs) was 1.02, with a 95% CI of 0.97–1.07 ( $p = 0.42$ ), thereby including the value of 1 and suggesting a high degree of concordance between the two assessments. Interestingly, 16 trials reported data on the disease control rate (DCR) with similar results to those observed for the ORR (Zhang et al., 2017). Consistently, in a subsequent large analysis of 76 Phase 3 clinical trials, including over 45,000 patients, the authors found eight trials with discrepancy in the evaluation of ORR among LIs and BICR. In seven of these trials, ORR was included as a secondary endpoint, further indicating that the ORR is rarely used as primary endpoint in Phase 3 clinical trials. Interestingly, in five out of eight trials, the BICR appeared to be more positive than the LI assessments (please refer to Table 2 in the original paper, Zhang et al. (2018)). The ORR was included as a primary endpoint in a clinical trial investigating the potential beneficial effect of peptide vaccine therapy as add-on treatment to aldesleukin (IL-2) in patients with locally advanced or metastatic melanoma (NCT00019682, Schwartzentruber et al., 2011). Interestingly, in this study, the BICR was more optimistic in the evaluation of the ORR in comparison to LI, although both evaluations found the addition of the vaccine more effective than the treatment with aldesleukin alone. Interestingly, the vaccine significantly increased the PFS and OS (although the study was not powered to assess these outcomes), indicating in this case a correlation between ORR and the other outcomes. However, the pooled analysis carried out on 29 trials that included the evaluation of ORR showed no significant differences among the two reviews (Zhang et al., 2018). Taken together, these data, therefore, suggest that there is an overall trend to overestimate the ORR by LIs. However, in controlled studies no significant differences can be found in the analysis of the beneficial effects of the investigational drug versus the comparator treatments among the two reviews, thus highlighting the reliability of the LIs in the evaluation of this endpoint in Phase 3 clinical trials.

### Progression-Free Survival

The PFS is defined as the time from randomization or treatment initiation until first evidence of objective tumor progression or death from any cause (Lebwohl et al., 2009; Amit et al., 2011; Delgado and Guddati, 2021). It is a time-to-event endpoint based on the assessment of disease progression at different time points defined in each clinical trial. It depends on tumor growth. This evaluation usually requires shorter time and a smaller population than OS, and it is not affected by subsequent therapies. Thus, it is often used as a primary endpoint in Phase 3 clinical trials, especially in the context of advanced/metastatic disease (Lebwohl et al., 2009). However, the data obtained may be subject to measurement error and bias since progression is defined based on standard criteria and is not objective as the time of death (Amit et al., 2010). That is why BICR is often

recommended to minimize possible investigator bias, except when the trial is truly double-blinded (Amit et al., 2010). However, also in these cases, the occurrence of adverse drug effects can be valuable information for the investigator to reveal the treatment arm in which the patient is possibly allocated (Dodd et al., 2008).

In an initial analysis of concordance among LIs and BICR, data from 7 Phase 3 trials showed that there were not significant differences in the assessment of PFS and treatment efficacy among the two reviews (Dodd et al., 2008). However, in a subsequent study, a certain degree of variability between the LI assessments and the BICR results with respect to the evaluation of PFS was found in eight Phase 3 clinical trials (Tang et al., 2010). This analysis focused mainly on differences in the median times, showing that the estimated mean difference in PFS was  $\geq 0.19$  (95% CI,  $-0.68$  to  $+0.29$ ) months. The concordance of treatment effect, as measured by the hazard ratio (HR) for LIs and BICR, could be analyzed only in four trials, two of which reported differences among the LIs and BICR (Tang et al., 2010). In contrast, a large meta-analysis, promoted by several pharmaceutical companies, namely, GlaxoSmithKline, AstraZeneca, Eli Lilly, and Hoffman LaRoche, and carried out on 27 Phase 3 trials with PFS, showed strong correlation ( $r = 0.947$ ) between the LI assessments and BICR. The estimated ratio of HRs between LIs and BICR was 1.03 (95% CI, 0.98–1.08), implying approximately 3% difference between the two evaluations (Amit et al., 2011). However, in a subsequent analysis of 76 Phase 3 clinical trials, 10 trials were identified with inconsistencies between the two reviews with respect to the assessment of PFS (Zhang et al., 2018). These studies mostly included PFS as the primary endpoint. In seven out of 10 trials, the LIs reported lower HR values in comparison to BICR, which means a larger beneficial effect of the investigational drug estimated locally. Consistent with data on the ORR discussed above, the pooled analysis carried out on 72 trials that included the evaluation of PFS showed no significant differences among the two reviews (please refer to Table 3 in the original paper, Zhang et al. (2018)). In a recent analysis performed by our group on 28 Phase 3 clinical trials, we compared the HRs for PFS between LIs and BICR calculating a discrepancy index through the ratio of the HRs among the two reviews, and we did not find any significant evaluation bias among the two assessments (Dello Russo et al., 2020). The estimated mean discrepancy index was 0.98 (95% CI, 0.927–1.032). We also performed subsequent analyses, dividing the trials by year, by tumor type, by drug type, and by study design. Interestingly, the discrepancy index did not include the value 1 only in the subgroup of trials started in the period 2003–2006, with the blinded assessment showing increased benefit than the LI assessment. We can hypothesize that the improvement in the diagnostic skills in clinical trials and the continuous refinement of diagnostic criteria per different tumors has progressively contributed to the reduction of evaluation bias. Consistently, we found a relatively higher variability (large CI) in trials testing biological agents, which was mainly due to the high

discrepancy recorded in trial KEYNOTE-61 testing the effects of the immune check point inhibitor pembrolizumab (Shitara et al., 2018). In this regard, it is now well established that standard RECIST criteria cannot be applied to the assessment of immunotherapy efficacy in solid tumors and needed to be modified for these drugs, so-called immune-related response evaluation criteria in solid tumors (iRECIST) (Seymour et al., 2017). No other relevant differences were found in the subgroup analysis based on tumor type (trials in onco-hematology and solid tumors) and drug type (trials testing small-molecules and those testing biologicals). We also found a trend toward more optimistic assessments by the independent review in double-blinded trials. Taken together, these data suggest that the two approaches, local and central review, with respect to PFS, lead to substantially similar results. These results further support the notion that the BICR is not necessary for all clinical trials, but it can be considered based on specific cases or carried out as an audit on a subset of enrolled patients (Amit et al., 2010).

## ADDITIONAL EVIDENCE ON THE COMPARISON BETWEEN LOCAL INVESTIGATORS AND INDEPENDENT REVIEWERS

In our previous analysis, we found that the LIs tend to overestimate the ORR in comparison to BICR, whereas high concordance was observed for the evaluation of PFS. This result was unexpected, considering that both endpoints rely on the same criteria to assess the radiological images as well as to establish response to therapy and/or disease progression. In an attempt to clarify the reasons behind the differences between LIs and BICR in relation to the evaluation of these two outcomes, we accessed the clinicaltrials.gov database and the clinicaltrialsregister.eu (EudraCT) database on 30 October 2021 and searched for clinical trials that included both the ORR and the PFS as endpoints and evaluations by both LIs and blinded assessors. We have already analyzed a subgroup of Phase 2 clinical trials that included both the ORR and the PFS, showing no differences between local and central reviewers with respect to PFS (Dello Russo et al., 2021). We aimed to update and extend this initial analysis by including also the Phase 3 clinical trials.

### Research Strategy

The research strategy was based on the methodological approach previously described (Dello Russo et al., 2020). In brief, in the clinicaltrials.gov database, we searched for clinical trials selected by the search string ((investigator-assessed) OR (investigator-assessment) AND (Cancer) and posted in the database within the data-lock date of 30 October 2021. We found a total of 1389 interventional clinical trials, of which 301 Phase 2 studies with results and 171 Phase 3 studies with results. Among these, we then selected for the analysis clinical trials that included both the ORR and the PFS evaluated both locally and by BICR. According to these inclusion criteria, we

selected 17 Phase 2 and 18 Phase 3 clinical trials. However, two Phase 2 clinical trials were subsequently excluded, namely, the NCT01258608 trial, because the result section only included ORR evaluated by BICR (64.4 months, 95% CI, 48.80–78.10) and the NCT02108964 clinical trial for which no results were posted in the database at the data-lock date.

A further search was carried out in the EudraCT database using the same search string as above, retrieving additional 614 clinical trials, 203 of with results including 116 Phase 2 and 88 Phase 3 clinical trials. Among these, we then selected for the analysis studies including both ORR and PFS evaluated both locally and by BICR. A total of nine Phase 2 and eight Phase 3 clinical trials met the inclusion criteria. Among the newly found Phase 2 clinical trials, only four did not overlap with the search in the clinicaltrials.gov database, although only two could be included for the analysis (Table 1, trials number 16 and 17). In this regard, the EudraCT 2013-000311-25 (NCT01915589, with no results) and the EudraCT 2013-000241-39 evaluating the efficacy of refametinib in the hepatocellular carcinoma were excluded since it was not possible to calculate a discrepancy index for the ORR. In addition, the Phase 2 clinical trial number 16 (Table 1), EudraCT 2013-005486-39, was linked to two different trials in the US database, namely, the NCT02108652 and NCT02951767 clinical trials, including results from two different cohorts of patients. Both trials are reported in Table 1 as number 16. For the newly found Phase 3 clinical trials, only three did not overlap with those found in the clinicaltrials.gov database. The total number of Phase 3 trials considered for the analysis is therefore 21.

### Results From Phase 2 Clinical Studies

As summarized in Table 1, a total of 17 clinical Phase 2 clinical trials were selected for the current analysis, including nine studies, highlighted in gray, considered in our previous analysis (Dello Russo et al., 2021). All the trials reported in Table 1 had an open-label design and mostly (10/17, 58.8%) a single group assignment. Among the remaining studies with parallel group assignment, only two included an active comparator group, that is, the NCT00875667 (n. 1) clinical trial (Trněný et al., 2016), studying the efficacy of lenalidomide *versus* chemotherapy as per investigator's choice in patients with mantle cell lymphoma after previous treatment failure, and the NCT01984242 (n. 5) clinical trial (McDermott et al., 2018), testing the efficacy of atezolizumab alone or in combination with bevacizumab *versus* sunitinib in patients affected by advanced renal carcinoma. All the other trials with parallel group assignment tested the same treatment in different subgroups or different schedule of administration of the same drug or drug combination. Therefore, all these studies substantially lacked a comparator arm. Interestingly, both the NCT00875667 and the NCT01984242 clinical trials, with the active comparator arm, included the PFS per BICR as primary endpoint. The latter was also found as the primary outcome in the NCT00117637 (n. 9) clinical trial (Escudier et al., 2009), testing the efficacy of sorafenib in combination with interferon- $\alpha$  (IFN- $\alpha$ ) in two different schedules of administration, that is, Cohort 1 (Sorafenib First) and Cohort 2 (IFN- $\alpha$  First). With respect to the

primary endpoint, the majority of the selected trials (7/17, 41%) included the evaluation of ORR per BICR, whereas only four (23.5%) trials reported the ORR per LI as the primary outcome. Interestingly, all these four trials were set to assess the efficacy of ALK1 inhibitors (three trials focused on ceritinib and one on crizotinib) in the context of non-small-cell lung cancer. In these trials, tumor assessment was performed every 8 weeks with a median time to first drug response of approximately of 2.0 months by BICR *versus* 2.3 months per LIs, which indicates a rapid pattern of tumor response to these inhibitors.

To evaluate the concordance between the two evaluations in this set of Phase 2 clinical trials, a discrepancy index (DI) was calculated for both the ORR (expressed as % of patients with CR and PR over the total number of the enrolled patients) and the PFS (considering the median PFS time for each treatment group) (Table 1). The DI was used in our previous analyses in order to assess differences between the two evaluations (Dello Russo et al., 2020; 2021). It is calculated as the ratio between the LI evaluation over the corresponding independently assessed endpoint, with a DI >1 indicating that the investigator was “more optimistic” and a DI <1 indicating the opposite, that the “blinded reviewer was more optimistic.” Among all these studies, the average DI for the ORR was 1.24 (95% CI, 1.005–1.478,  $n = 31$ ) and the average DI for PFS was 1.05 (95% CI, 0.958–1.141). These data suggest that the LIs tend to overestimate the ORR, with a +24% of discrepancy which is little over what observed in our previous analysis (average +17.5% of positive responses, Dello Russo et al., 2021). In line with previous data, we found a substantial agreement between the two evaluations for the PFS in these Phase 2 clinical trials. Interestingly, in the abovementioned four clinical trials on ALK inhibitors having the ORR per LI as the primary outcome, we found an average DI of 1.19 for both the ORR and the PFS which suggest a +19% overestimation of positive effects by LIs in comparison to BICR on both endpoints.

In Table 2, we summarized the results for eight different trials data on the time to first treatment response (TTFR), the time to progression (TTP) and the median PFS. From the comparison of these different endpoints, it emerges that the average time to observe a tumor response to treatment is relatively shorter than the time to reach progression. This is an expected finding, which implies that a reduced number of scans are indeed necessary to observe the response to treatment, therefore explaining in part the higher variability observed in the assessment of ORR locally and centrally. With time and the evaluation of multiple scans, the variability among the two evaluations tends to be reduced. Moreover, we can also hypothesize that it is more difficult to detect a response to treatment particularly in term of distinction between PR and stable disease in comparison to disease progression, which may further contribute to the higher discrepancy observed for the ORR among the two evaluations.

## Results From Phase 3 Clinical Studies

A total of 21 Phase 3 clinical trials, including both ORR and PFS evaluated locally and centrally were found (Table 3). These included one trial, the NCT00075270 ( $n = 3$ ) (Di Leo

et al., 2008), in which the time to disease progression (TTP) per LI and BRIC was included as primary endpoint and considered in place of the PFS for the present evaluation. The trial NCT01287741 ( $n = 15$ ) only included the HR for the evaluation of PFS by LI and BICR (Vitolo et al., 2017). Both trials were also included in our previous analysis (Dello Russo et al., 2020), therefore kept in the present analysis. The majority of these studies (17/21, 81%) had an open-label design and (10/21, 47.6%) included the evaluation of PFS per BICR as primary endpoint. In this regard, only six studies out of 21 (28.6%) included the evaluation of PFS per LI as the primary outcome, whereas three studies (14.2%) had both evaluations as the primary endpoint.

To assess the concordance between the two evaluations in this set of Phase 3 clinical trials, a discrepancy index (DI) was calculated for both the ORR (expressed as % of patients with CR and PR over the total of the enrolled patients) and the PFS (considering the median PFS time for each treatment group). Among all these studies, the average DI for the ORR was 1.20 (95% CI, 1.075–1.328,  $n = 42$ ) and for PFS was 1.014 (95% CI, 0.963–1.065,  $n = 42$ ). These data suggest that the LIs tend to overestimate the ORR, by a factor of +20% in the selected Phase 3 clinical trials, whereas there is a substantial agreement between the two evaluations for the PFS when considering the PFS median time.

In addition, for the ORR, we calculated the DI for the odd ratios (ORs), when reported (four trials), and for the PFS, the DI based on the ratio of HRs was calculated per type of evaluation for most of the trials. The DI calculated on the ORs for the ORR was 1.04, 95% CI 0.912–1.168, indicating a potential agreement when it comes to the evaluation of treatment effect among the two evaluations. With respect of the DI calculated on the HRs for PFS, out of 21 trials included in the present analysis, 13 were already considered in our previous study (Dello Russo et al., 2020), whereas six trials brought additional new data. On average, the DI calculated on the HRs from these six trials was lower than that from the previous observations (0.84, 95% CI 0.756–0.923,  $n = 6$ ). However, by pooling these data with those of previous analysis, the average DI value was only slightly reduced, and 95% CI included the value 1 (0.958, 95% CI 0.91–1.01,  $n = 38$ ).

As shown in Table 4, for three Phase 3 trials, it was possible to evaluate the TTFR together with the PFS. In line with data from Phase 2 clinical trials, the TTFR was shorter than the median PFS, with minimal differences observed in NCT02603432 ( $n = 17$ ) trial evaluating the efficacy of avelumab in combination with best supportive care as maintenance therapy in patients with locally advanced or metastatic urothelial cancer with stable disease after completion of first-line platinum-containing chemotherapy (Powles et al., 2020). However, in this study only 34 responses out of 350 treated patients were observed in the experimental group in comparison to five responses out of 350 patients in the comparator arm by BICR. The LI reported a higher number of tumor responses in comparison to BICR. The fact that the ORR is often calculated based on a reduced number of responses may also contribute to the higher variability observed for the assessment of this endpoint by LIs and BICR.

## CONCLUSION

Considering the review of available evidence carried out in this study, along with additional new evidence shown here for the first time, we can attempt to draw some overall conclusions:

- ✓ The additional analysis presented in this work confirmed that 1) LIs tend to overestimate the ORR in comparison to BICR, whereas 2) no significant differences are observed between LI and BICR concerning the assessment of PFS.
- ✓ The assessment of ORR is endowed with higher variability than that of PFS. Such larger variability is associated with various factors, including 1) a limited number of measurements, compared to repeated measures with PFS; 2) the time-to-response, which is a variable in ORR assessment, whereas PFS is always measured once a response is established; 3) the time-to-response is in turn influenced by the type of treatment, with small molecules in general inducing faster responses than immunotherapies; 4) some protocols may assess ORR at fixed times, while other may consider the *best response* to calculate ORR. Thus, a higher variability seems to be associated to a significant expectation bias.
- ✓ The analysis of ORR assessment in Phase 3 trials showed that, luckily, LIs tend to overestimate ORR compared to BICR both in experimental and control groups. Thus, by analyzing the data as ORs, the overestimations of control ORRs tend to counterbalance those of experimental ORRs, thereby reducing the gap between LI and BICR.

## REFERENCES

- Amit, O., Bushnell, W., Dodd, L., Roach, N., and Sargent, D. (2010). Blinded Independent central Review of the Progression-free Survival Endpoint. *Oncologist* 15 (5), 492–495. doi:10.1634/theoncologist.2009-0261
- Amit, O., Mannino, F., Stone, A. M., Bushnell, W., Denne, J., Helterbrand, J., et al. (2011/1990). Blinded Independent central Review of Progression in Cancer Clinical Trials: Results from a Meta-Analysis. *Eur. J. Cancer* 47 (12), 1772–1778. doi:10.1016/j.ejca.2011.02.013
- Armand, P., Murawski, N., Molin, D., Zain, J., Eichhorst, B., Gulbas, Z., et al. (2020). Pembrolizumab in Relapsed or Refractory Richter Syndrome. *Br. J. Haematol.* 190 (2), e117–e120. doi:10.1111/bjh.16762
- Armand, P., Rodig, S., Melnichenko, V., Thieblemont, C., Bouabdallah, K., Tulyan, G., et al. (2019). Pembrolizumab in Relapsed or Refractory Primary Mediastinal Large B-Cell Lymphoma. *J. Clin. Oncol.* 37 (34), 3291–3299. doi:10.1200/JCO.19.01389
- Aykan, N. F., and Özatlı, T. (2020). Objective Response Rate Assessment in Oncology: Current Situation and Future Expectations. *World J. Clin. Oncol.* 11 (2), 53–73. doi:10.5306/wjco.v11.i2.53
- Balar, A. V., Galsky, M. D., Rosenberg, J. E., Powles, T., Petrylak, D. P., Bellmunt, J., et al. IMvigor210 Study Group (2017). Atezolizumab as First-Line Treatment in Cisplatin-Ineligible Patients with Locally Advanced and Metastatic Urothelial Carcinoma: a Single-Arm, Multicentre, Phase 2 Trial. *Lancet* 389 (10064), 67–76. doi:10.1016/S0140-6736(16)32455-2
- Benboubker, L., Dimopoulos, M. A., Dispenzieri, A., Catalano, J., Belch, A. R., Cavo, M., et al. and FIRST Trial Team (2014). Lenalidomide and Dexamethasone in Transplant-Ineligible Patients with Myeloma. *N. Engl. J. Med.* 371 (10), 906–917. doi:10.1056/NEJMoa1402551
- Chen, E. Y., Raghunathan, V., and Prasad, V. (2019). An Overview of Cancer Drugs Approved by the US Food and Drug Administration Based on the Surrogate End Point of Response Rate. *JAMA Intern. Med.* 179 (7), 915–921. doi:10.1001/jamainternmed.2019.0583

## FUNDING

No specific funding was allocated for this study.

## AUTHOR CONTRIBUTIONS

CDR carried out the literature review and extracted the data from the clinicaltrials.gov database. CDR and PN carried out the analysis on the additional trials. CDR wrote the first draft of the manuscript. PN contributed to the design, writing, and revision of the manuscript. Both authors have approved the submitted version.

## ACKNOWLEDGMENTS

The authors would like to thank Prof. Emilio Bria, Medical Oncology, Comprehensive Cancer Center, Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Università Cattolica del Sacro Cuore, for meaningful discussion on the relevance of clinical endpoints in oncology drug development. The authors would also like to acknowledge the initial work on the clinicaltrials.gov database carried out by Dr Natalia Cappoli, a PhD student in Neuroscience, Università Cattolica del Sacro Cuore, and Dr Daniela Pilunni, Postgraduate School of Hospital Pharmacy, Sapienza University, Rome, Italy.

- Chow, L. Q., Barlesi, F., Bertino, E. M., van den Bent, M. J., Wakelee, H., Wen, P. Y., et al. (2019). Results of the ASCEND-7 Phase II Study Evaluating ALK Inhibitor (ALKi) Ceritinib in Patients (Pts) with ALK+ Non-small Cell Lung Cancer (NSCLC) Metastatic to the Brain. *Ann. Oncol.* 30 (5 Suppl. 1), v602–v603. doi:10.1093/annonc/mdz260
- Clinicaltrials.gov identifier NCT01708174 (2012). Available at: <https://clinicaltrials.gov/ct2/show/NCT01708174>. Published 2012. Accessed 7/01/2022.
- Cooper, K., Tappenden, P., Cantrell, A., and Ennis, K. (2020). A Systematic Review of Meta-Analyses Assessing the Validity of Tumour Response Endpoints as Surrogates for Progression-free or Overall Survival in Cancer. *Br. J. Cancer* 123 (11), 1686–1696. doi:10.1038/s41416-020-01050-w
- Daniele, G., Giannarelli, D., and Bria, E. (2020). Looking for a Better Measure of the Benefit in Clinical Trials: a Never-Ending Journey. *Ann. Transl. Med.* 8 (14), 893. doi:10.21037/atm.2020.03.118
- Delgado, A., and Guddati, A. K. (2021). Clinical Endpoints in Oncology - a Primer. *Am. J. Cancer Res.* 11 (4), 1121–1131.
- Dello Russo, C., Cappoli, N., and Navarra, P. (2020). A Comparison between the Assessments of Progression-free Survival by Local Investigators versus Blinded Independent central Reviews in Phase III Oncology Trials. *Eur. J. Clin. Pharmacol.* 76 (8), 1083–1092. doi:10.1007/s00228-020-02895-z
- Dello Russo, C., Cappoli, N., Pilunni, D., and Navarra, P. (2021). Local Investigators Significantly Overestimate Overall Response Rates Compared to Blinded Independent Central Reviews in Phase 2 Oncology Trials. *J. Clin. Pharmacol.* 61 (6), 810–819. doi:10.1002/jcph.1790
- Di Leo, A., Gomez, H. L., Aziz, Z., Zvirbule, Z., Bines, J., Arbushites, M. C., et al. (2008). Phase III, Double-Blind, Randomized Study Comparing Lapatinib Plus Paclitaxel with Placebo Plus Paclitaxel as First-Line Treatment for Metastatic Breast Cancer. *J. Clin. Oncol.* 26 (34), 5544–5552. doi:10.1200/JCO.2008.16.2578
- Dodd, L. E., Korn, E. L., Freidlin, B., Jaffe, C. C., Rubinstein, L. V., Dancey, J., et al. (2008). Blinded Independent Central Review of Progression-free Survival in Phase III Clinical Trials: Important Design Element or Unnecessary Expense? *J. Clin. Oncol.* 26 (22), 3791–3796. doi:10.1200/JCO.2008.16.1711

- Dreyling, M., Morschhauser, F., Bouabdallah, K., Bron, D., Cunningham, D., Assouline, S. E., et al. (2017). Phase II Study of Copanlisib, a PI3K Inhibitor, in Relapsed or Refractory, Indolent or Aggressive Lymphoma. *Ann. Oncol.* 28 (9), 2169–2178. doi:10.1093/annonc/mdx289
- Eisenhauer, E. A., Therasse, P., Bogaerts, J., Schwartz, L. H., Sargent, D., Ford, R., et al. (2009). New Response Evaluation Criteria in Solid Tumours: Revised RECIST Guideline (Version 1.1). *Eur. J. Cancer*, 45 (2), 228–247. doi:10.1016/j.ejca.2008.10.026
- Escudier, B., Eisen, T., Stadler, W. M., Szczylik, C., Oudard, S., Siebels, M., et al. (2007). Sorafenib in Advanced Clear-cell Renal-Cell Carcinoma. *N. Engl. J. Med.* 356 (2), 125–134. doi:10.1056/NEJMoa060655
- Escudier, B., Szczylik, C., Hutson, T. E., Demkow, T., Staehler, M., Rolland, F., et al. (2009). Randomized Phase II Trial of First-Line Treatment with Sorafenib versus Interferon Alfa-2a in Patients with Metastatic Renal Cell Carcinoma. *J. Clin. Oncol.* 27 (8), 1280–1289. doi:10.1200/JCO.2008.19.3342
- Flaherty, K. T., Infante, J. R., Daud, A., Gonzalez, R., Kefford, R. F., Sosman, J., et al. (2012). Combined BRAF and MEK Inhibition in Melanoma with BRAF V600 Mutations. *N. Engl. J. Med.* 367 (18), 1694–1703. doi:10.1056/NEJMoa1210093
- Ford, R., Schwartz, L., Dancy, J., Dodd, L. E., Eisenhauer, E. A., Gwyther, S., et al. (2009). Lessons Learned from Independent Central Review. *Eur. J. Cancer* 45 (2), 268–274. doi:10.1016/j.ejca.2008.10.031
- Gaillard, S., Oaknin, A., Ray-Coquard, I., Vergote, I., Scambia, G., Colombo, N., et al. (2021). Lurbinectedin versus Pegylated Liposomal Doxorubicin or Topotecan in Patients with Platinum-Resistant Ovarian Cancer: A Multicenter, Randomized, Controlled, Open-Label Phase 3 Study (CORAIL). *Gynecol. Oncol.* 163 (2), 237–245. doi:10.1016/j.ygyno.2021.08.032
- Global Cancer Observatory (2020). *Cancer Today*. International Agency for Research on Cancer. (Accessed October 24, 2021).
- Gogas, H., Dréno, B., Larkin, J., Demidov, L., Stroyakovskiy, D., Eroglu, Z., et al. (2021). Cobimetinib Plus Atezolizumab in BRAFV600 Wild-type Melanoma: Primary Results from the Randomized Phase III IMspire170 Study. *Ann. Oncol.* 32 (3), 384–394. doi:10.1016/j.annonc.2020.12.004
- Gravanis, I., Vleminckx, C., Jonsson, B., and Pignatti, F. (2014). The Changing World of Cancer Drug Development: The Regulatory Bodies' Perspective. *Chin. Clin. Oncol.* 3 (2), 22. doi:10.3978/j.issn.2304-3865.2014.05.08
- Haslam, A., Hey, S. P., Gill, J., and Prasad, V. (2019). A Systematic Review of Trial-Level Meta-Analyses Measuring the Strength of Association Between Surrogate End-Points and Overall Survival in Oncology. *Eur. J. Cancer* 106, 196–211. doi:10.1016/j.ejca.2018.11.012
- Hida, T., Satouchi, M., Nakagawa, K., Seto, T., Matsumoto, S., Kiura, K., et al. (2017). Ceritinib in Patients with Advanced, Crizotinib-Treated, Anaplastic Lymphoma Kinase-Rearranged NSCLC: Japanese Subset. *Jpn. J. Clin. Oncol.* 47 (7), 618–624. doi:10.1093/jcco/hyx045
- Hiddemann, W., Barbui, A. M., Canales, M. A., Cannell, P. K., Collins, G. P., Dürig, J., et al. (2018). Immunochemotherapy with Obinutuzumab or Rituximab for Previously Untreated Follicular Lymphoma in the GALLIUM Study: Influence of Chemotherapy on Efficacy and Safety. *J. Clin. Oncol.* 36 (23), 2395–2404. doi:10.1200/JCO.2017.76.8960
- Iannesi, A., Beaumont, H., Liu, Y., and Bertrand, A. S. (2021). RECIST 1.1 and Lesion Selection: How to deal with Ambiguity at Baseline? *Insights Imaging* 12 (1), 36. doi:10.1186/s13244-021-00976-w
- Ikedo, M., Ioka, T., Fukutomi, A., Morizane, C., Kasuga, A., Takahashi, H., et al. (2018). Efficacy and Safety of Trametinib in Japanese Patients with Advanced Biliary Tract Cancers Refractory to Gemcitabine. *Cancer Sci.* 109 (1), 215–224. doi:10.1111/cas.13438
- Kim, D. W., Tiseo, M., Ahn, M. J., Reckamp, K. L., Hansen, K. H., Kim, S. W., et al. (2017). Brigatinib in Patients with Crizotinib-Refractory Anaplastic Lymphoma Kinase-Positive Non-small-cell Lung Cancer: A Randomized, Multicenter Phase II Trial. *J. Clin. Oncol.* 35 (22), 2490–2498. doi:10.1200/JCO.2016.71.5904
- Krop, I. E., LoRusso, P., Miller, K. D., Modi, S., Yardley, D., Rodriguez, G., et al. (2012). A Phase II Study of Trastuzumab Emtansine in Patients with Human Epidermal Growth Factor Receptor 2-positive Metastatic Breast Cancer Who Were Previously Treated with Trastuzumab, Lapatinib, an Anthracycline, a Taxane, and Capecitabine. *J. Clin. Oncol.* 30 (26), 3234–3241. doi:10.1200/JCO.2011.40.5902
- Kwong, G. A., Ghosh, S., Gamboa, L., Patriotis, C., Srivastava, S., and Bhatia, S. N. (2021). Synthetic Biomarkers: a Twenty-First Century Path to Early Cancer Detection. *Nat. Rev. Cancer* 21 (10), 655–668. doi:10.1038/s41568-021-00389-3
- Lebwohl, D., Kay, A., Berg, W., Baladi, J. F., and Zheng, J. (2009). Progression-free Survival: Gaining on Overall Survival as a Gold Standard and Accelerating Drug Development. *Cancer J.* 15 (5), 386–394. doi:10.1097/PPO.0b013e3181b9c5ec
- Lonial, S., Lee, H. C., Badros, A., Trudel, S., Nooka, A. K., Chari, A., et al. (2020). Belantamab Mafodotin for Relapsed or Refractory Multiple Myeloma (DREAMM-2): a Two-Arm, Randomised, Open-Label, Phase 2 Study. *Lancet Oncol.* 21 (2), 207–221. doi:10.1016/S1470-2045(19)30788-0
- Marcus, R., Davies, A., Ando, K., Klapper, W., Opat, S., Owen, C., et al. (2017). Obinutuzumab for the First-Line Treatment of Follicular Lymphoma. *N. Engl. J. Med.* 377 (26), 1331–1344. doi:10.1056/NEJMoa1614598
- Matulonis, U. A., Shapira-Frommer, R., Santin, A. D., Lisianskaya, A. S., Pignata, S., Vergote, I., et al. (2019). Antitumor Activity and Safety of Pembrolizumab in Patients with Advanced Recurrent Ovarian Cancer: Results from the Phase II KEYNOTE-100 Study. *Ann. Oncol.* 30 (7), 1080–1087. doi:10.1093/annonc/mdz135
- McDermott, D. F., Huseni, M. A., Atkins, M. B., Motzer, R. J., Rini, B. I., Escudier, B., et al. (2018). Clinical Activity and Molecular Correlates of Response to Atezolizumab Alone or in Combination with Bevacizumab versus Sunitinib in Renal Cell Carcinoma. *Nat. Med.* 24 (6), 749–757. doi:10.1038/s41591-018-0053-3
- Michels, S., Massutí, B., Schildhaus, H. U., Franklin, J., Sebastian, M., Felip, E., et al. (2019). Safety and Efficacy of Crizotinib in Patients with Advanced or Metastatic ROS1-Rearranged Lung Cancer (EUCROSS): A European Phase II Clinical Trial. *J. Thorac. Oncol.* 14 (7), 1266–1276. doi:10.1016/j.jtho.2019.03.020
- Monk, B. J., Colombo, N., Oza, A. M., Fujiwara, K., Birrer, M. J., Randall, L., et al. (2021). Chemotherapy with or without Avelumab Followed by Avelumab Maintenance versus Chemotherapy Alone in Patients with Previously Untreated Epithelial Ovarian Cancer (JAVELIN Ovarian 100): An Open-Label, Randomised, Phase 3 Trial. *Lancet Oncol.* 22 (9), 1275–1289. doi:10.1016/S1470-2045(21)00342-9
- Motzer, R. J., Hutson, T. E., Tomczak, P., Michaelson, M. D., Bukowski, R. M., Rixe, O., et al. (2007). Sunitinib versus Interferon Alfa in Metastatic Renal-Cell Carcinoma. *N. Engl. J. Med.* 356 (2), 115–124. doi:10.1056/NEJMoa065044
- Nishio, M., Felip, E., Orlov, S., Park, K., Yu, C. J., Tsai, C. M., et al. (2020). Final Overall Survival and Other Efficacy and Safety Results from ASCEND-3: Phase II Study of Ceritinib in ALKI-Naive Patients with ALK-Rearranged NSCLC. *J. Thorac. Oncol.* 15 (4), 609–617. doi:10.1016/j.jtho.2019.11.006
- Novello, S., Mazières, J., Oh, I. J., de Castro, J., Migliorino, M. R., Helland, Å., et al. (2018). Alectinib versus Chemotherapy in Crizotinib-Pretreated Anaplastic Lymphoma Kinase (ALK)-positive Non-small-cell Lung Cancer: Results from the Phase III ALUR Study. *Ann. Oncol.* 29 (6), 1409–1416. doi:10.1093/annonc/mdy121
- Okoli, G. N., Lam, O. L. T., Reddy, V. K., Copstein, L., Askin, N., Prashad, A., et al. (2021). Interventions to Improve Early Cancer Diagnosis of Symptomatic Individuals: A Scoping Review. *BMJ open* 11 (11), e055488. doi:10.1136/bmjopen-2021-055488
- Panayiotidis, P., Follows, G. A., Mollica, L., Nagler, A., Özcan, M., Santoro, A., et al. (2021). Efficacy and Safety of Copanlisib in Patients with Relapsed or Refractory Marginal Zone Lymphoma. *Blood Adv.* 5 (3), 823–828. doi:10.1182/bloodadvances.2020002910
- Pasalic, D., McGinnis, G. J., Fuller, C. D., Grossberg, A. J., Verma, V., Mainwaring, W., et al. (2020). Progression-free Survival Is a Suboptimal Predictor for Overall Survival Among Metastatic Solid Tumour Clinical Trials. *Eur. J. Cancer* 136, 176–185. doi:10.1016/j.ejca.2020.06.015
- Paz-Ares, L., de Marinis, F., Dediu, M., Thomas, M., Pujol, J. L., Bidoli, P., et al. (2012). Maintenance Therapy with Pemetrexed Plus Best Supportive Care versus Placebo Plus Best Supportive Care after Induction Therapy with Pemetrexed Plus Cisplatin for Advanced Non-squamous Non-small-cell Lung Cancer (PARAMOUNT): a Double-Blind, Phase 3, Randomised Controlled Trial. *Lancet Oncol.* 13 (3), 247–255. doi:10.1016/S1470-2045(12)70063-3
- Perez, E. A., Barrios, C., Eiermann, W., Toi, M., Im, Y. H., Conte, P., et al. (2017). Trastuzumab Emtansine with or without Pertuzumab versus Trastuzumab Plus Taxane for Human Epidermal Growth Factor Receptor 2-Positive, Advanced Breast Cancer: Primary Results from the Phase III Marianne Study. *J. Clin. Oncol.* 35 (2), 141–148. doi:10.1200/JCO.2016.67.4887
- Perez-Gracia, J. L., Loriot, Y., Rosenberg, J. E., Powles, T., Necchi, A., Hussain, S. A., et al. (2018). Atezolizumab in Platinum-Treated Locally Advanced or Metastatic Urothelial Carcinoma: Outcomes by Prior Number of Regimens. *Eur. Urol.* 73 (3), 462–468. doi:10.1016/j.eururo.2017.11.023

- Peters, S., Gettinger, S., Johnson, M. L., Jänne, P. A., Garassino, M. C., Christoph, D., et al. (2017). Phase II Trial of Atezolizumab as First-Line or Subsequent Therapy for Patients with Programmed Death-Ligand 1-Selected Advanced Non-small-cell Lung Cancer (BIRCH). *J. Clin. Oncol.* 35 (24), 2781–2789. doi:10.1200/JCO.2016.71.9476
- Powles, T., Park, S. H., Voog, E., Caserta, C., Valderrama, B. P., Gurney, H., et al. (2020). Avelumab Maintenance Therapy for Advanced or Metastatic Urothelial Carcinoma. *N. Engl. J. Med.* 383 (25), 1218–1230. doi:10.1056/NEJMoa2002788
- Pujade-Lauraine, E., Fujiwara, K., Ledermann, J. A., Oza, A. M., Kristeleit, R., Ray-Coquard, I. L., et al. (2021). Avelumab Alone or in Combination with Chemotherapy versus Chemotherapy Alone in Platinum-Resistant or Platinum-Refractory Ovarian Cancer (JAVELIN Ovarian 200): an Open-Label, Three-Arm, Randomised, Phase 3 Study. *Lancet Oncol.* 22 (7), 1034–1046. doi:10.1016/S1470-2045(21)00216-3
- Ramalingam, S. S., Jänne, P. A., Mok, T., O'Byrne, K., Boyer, M. J., Von Pawel, J., et al. (2014). Dacomitinib versus Erlotinib in Patients with Advanced-Stage, Previously Treated Non-small-cell Lung Cancer (ARCHER 1009): a Randomised, Double-Blind, Phase 3 Trial. *Lancet Oncol.* 15 (12), 1369–1378. doi:10.1016/S1470-2045(14)70452-8
- Rini, B. I., Powles, T., Atkins, M. B., Escudier, B., McDermott, D. F., Suarez, C., et al. IMmotion151 Study Group (2019). Atezolizumab Plus Bevacizumab versus Sunitinib in Patients with Previously Untreated Metastatic Renal Cell Carcinoma (IMmotion151): a Multicentre, Open-Label, Phase 3, Randomised Controlled Trial. *Lancet* 393 (10189), 2404–2415. doi:10.1016/S0140-6736(19)30723-8
- Robinson, A. G., Booth, C. M., and Eisenhauer, E. A. (2014). Progression-free Survival as an End-point in Solid Tumours—Perspectives from Clinical Trials and Clinical Practice. *Eur. J. Cancer* 50 (13), 2303–2308. doi:10.1016/j.ejca.2014.05.024
- Rothenberg, M. L., Cox, J. V., Butts, C., Navarro, M., Bang, Y. J., Goel, R., et al. (2008). Capecitabine Plus Oxaliplatin (XELOX) versus 5-fluorouracil/folinic Acid Plus Oxaliplatin (FOLFOX-4) as Second-Line Therapy in Metastatic Colorectal Cancer: a Randomized Phase III Noninferiority Study. *Ann. Oncol.* 19 (10), 1720–1726. doi:10.1093/annonc/mdn370
- Schwartzentruber, D. J., Lawson, D. H., Richards, J. M., Conry, R. M., Miller, D. M., Treisman, J., et al. (2011). gp100 Peptide Vaccine and Interleukin-2 in Patients with Advanced Melanoma. *N. Engl. J. Med.* 364 (22), 2119–2127. doi:10.1056/NEJMoa1012863
- Seymour, L., Bogaerts, J., Perrone, A., Ford, R., Schwartz, L. H., Mandrekar, S., et al. RECIST working group (2017). iRECIST: Guidelines for Response Criteria for Use in Trials Testing Immunotherapeutics. *Lancet Oncol.* 18 (3), e143–e152. doi:10.1016/S1470-2045(17)30074-8
- Shaw, A. T., Bauer, T. M., de Marinis, F., Felip, E., Goto, Y., Liu, G., et al. (2020). CROWN Trial Investigators First-Line Lorlatinib or Crizotinib in Advanced ALK-Positive Lung Cancer. *N. Engl. J. Med.* 383 (21), 2018–2029. doi:10.1056/nejmoa2027187
- Shitara, K., Özgüroğlu, M., Bang, Y. J., Di Bartolomeo, M., Mandalà, M., Ryu, M. H., Fornaro, L., Olesiński, T., Caglevic, C., Chung, H. C., Muro, K., Goekkurt, E., Mansoor, W., McDermott, R. S., Shacham-Shmueli, E., Chen, X., Mayo, C., Kang, S. P., Ohtsu, A., and Fuchs, C. S. (2018). Pembrolizumab versus Paclitaxel for Previously Treated, Advanced Gastric or Gastro-Oesophageal Junction Cancer (KEYNOTE-061): A Randomised, Open-Label, Controlled, Phase 3 Trial. *Lancet* 392, 123–133. doi:10.1016/S0140-6736(18)31257-1
- Spicka, I., Ocio, E. M., Oakervee, H. E., Greil, R., Banh, R. H., Huang, S. Y., et al. (2019). Randomized Phase III Study (ADMYRE) of Plitidepsin in Combination with Dexamethasone vs. Dexamethasone Alone in Patients with Relapsed/Refractory Multiple Myeloma. *Ann. Hematol.* 98 (9), 2139–2150. doi:10.1007/s00277-019-03739-2
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A. Cancer J. Clin.* 71 (3), 209–249. doi:10.3322/caac.21660
- Tang, P. A., Pond, G. R., and Chen, E. X. (2010). Influence of an Independent Review Committee on Assessment of Response Rate and Progression-free Survival in Phase III Clinical Trials. *Ann. Oncol.* 21 (1), 19–26. doi:10.1093/annonc/mdp478
- Therasse, P., Arbuck, S. G., Eisenhauer, E. A., Wanders, J., Kaplan, R. S., Rubinstein, L., et al. (2000). New Guidelines to Evaluate the Response to Treatment in Solid Tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J. Natl. Cancer Inst.* 92 (3), 205–216. doi:10.1093/jnci/92.3.205
- Trněný, M., Lamy, T., Walewski, J., Belada, D., Mayer, J., Radford, J., et al. (2016). Lenalidomide versus Investigator's Choice in Relapsed or Refractory Mantle Cell Lymphoma (MCL-002; SPRINT): a Phase 2, Randomised, Multicentre Trial. *Lancet Oncol.* 17 (3), 319–331. doi:10.1016/S1470-2045(15)00559-8
- Vitolo, U., Trněný, M., Belada, D., Burke, J. M., Carella, A. M., Chua, N., et al. (2017). Obinutuzumab or Rituximab Plus Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone in Previously Untreated Diffuse Large B-Cell Lymphoma. *J. Clin. Oncol.* 35 (31), 3529–3537. doi:10.1200/JCO.2017.73.3402
- Wu, Y.-L., Shi, Y., Tan, D. S. W., Xiaoqing, L., Cheng, Y., Zhou, J., et al. (2020). Phase 1/2 Study of Ceritinib in Chinese Patients with Advanced Anaplastic Lymphoma Kinase-Rearranged Non-small Cell Lung Cancer Previously Treated with Crizotinib: Results from ASCEND-6. *Lung Cancer* 150, 240–246. doi:10.1016/j.lungcan.2020.10.024
- Wu, Y. L., Cheng, Y., Zhou, X., Lee, K. H., Nakagawa, K., Niho, S., et al. (2017). Dacomitinib versus Gefitinib as First-Line Treatment for Patients with EGFR-Mutation-Positive Non-small-cell Lung Cancer (ARCHER 1050): a Randomised, Open-Label, Phase 3 Trial. *Lancet Oncol.* 18 (11), 1454–1466. doi:10.1016/S1470-2045(17)30608-3
- Zhang, J., Zhang, Y., Tang, S., Jiang, L., He, Q., Hamblin, L. T., et al. (2018). Systematic Bias Between Blinded Independent Central Review and Local Assessment: Literature Review and Analyses of 76 Phase III Randomised Controlled Trials in 45 688 Patients with Advanced Solid Tumour. *BMJ Open* 8 (9), e017240. doi:10.1136/bmjopen-2017-017240
- Zhang, J., Zhang, Y., Tang, S., Liang, H., Chen, D., Jiang, L., et al. (2017). Evaluation Bias in Objective Response Rate and Disease Control Rate Between Blinded Independent Central Review and Local Assessment: A Study-Level Pooled Analysis of Phase III Randomized Control Trials in the Past Seven Years. *Ann. Transl. Med.* 5 (24), 481. doi:10.21037/atm.2017.11.24

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Dello Russo and Navarra. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.