Design characteristics, risk of bias, and reporting of randomised controlled trials supporting approvals of cancer drugs by European Medicines Agency, 2014-16: cross sectional analysis

Huseyin Naci,[1,2] Courtney Davis,[3] Jelena Savović,[4,5] Julian P T Higgins,[4,5,6] Jonathan A C Sterne,[4,6] Bishal Gyawali,[2,7] Xochitl Romo-Sandoval,[1] Nicola Handley,[3] Christopher M Booth[7]

For numbered affiliations see end of the article.

Correspondence to: H Naci
h.naci@lse.ac.uk
(or @huseyinnaci2 on Twitter
ORCID 0000-0002-7192-5751)

ABSTRACT

OBJECTIVE
To examine the design characteristics, risk of bias, and reporting adequacy of pivotal randomised controlled trials of cancer drugs approved by the European Medicines Agency (EMA).

DESIGN
Cross sectional analysis.

SETTING
European regulatory documents, clinical trial registry records, protocols, journal publications, and supplementary appendices.

ELIGIBILITY CRITERIA
Pivotal randomised controlled trials of new cancer drugs approved by the EMA between 2014 and 2016.

MAIN OUTCOME MEASURES
Study design characteristics (randomisation, comparators, and endpoints); risk of bias using the revised Cochrane tool (bias arising from the randomisation process, deviations from intended interventions, missing outcome data, measurement of the outcome, and selection of the reported result); and reporting adequacy (completeness and consistency of information in trial protocols, publications, supplementary appendices, clinical trial registry records, and regulatory documents).

RESULTS
Between 2014 and 2016, the EMA approved 32 new cancer drugs on the basis of 54 pivotal studies. Of these, 41 (76%) were randomised controlled trials and 13 (24%) were either non-randomised studies or single arm studies. 39/41 randomised controlled trials had available publications and were included in our study. Only 10 randomised controlled trials (26%) measured overall survival as either a primary or coprimary endpoint, with the remaining trials evaluating surrogate measures such as progression free survival and response rates. Overall, 19 randomised controlled trials (49%) were judged to be at high risk of bias for their primary outcome. Concerns about missing outcome data (n=10) and measurement of the outcome (n=7) were the most common domains leading to high risk of bias judgments. Fewer randomised controlled trials that evaluated overall survival as the primary endpoint were at high risk of bias than those that evaluated surrogate efficacy endpoints (2/10 (20%) *v* 16/29 (55%), respectively). When information available in regulatory documents and the scientific literature was considered separately, overall risk of bias judgments differed for eight randomised controlled trials (21%), which reflects reporting inadequacies in both sources of information. Regulators identified additional deficits beyond the domains captured in risk of bias assessments for 10 drugs (31%). These deficits included magnitude of clinical benefit, inappropriate comparators, and non-preferred study endpoints, which were not disclosed as limitations in scientific publications.

CONCLUSIONS
Most pivotal studies forming the basis of EMA approval of new cancer drugs between 2014 and 2016 were randomised controlled trials. However, almost half of these were judged to be at high risk of bias based on their design, conduct, or analysis, some of which might be unavoidable because of the complexity of cancer trials. Regulatory documents and the scientific literature had gaps in their reporting. Journal publications did not acknowledge the key limitations of the available evidence identified in regulatory documents.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Most new cancer drugs approved by the European Medicines Agency (EMA) have been studied in randomised controlled trials, considered to be the "gold standard" for evaluating treatment efficacy

Design characteristics, risk of bias, and reporting of randomised controlled trials that support recent EMA cancer drug approvals have not been evaluated

WHAT THIS STUDY ADDS

Around half of randomised controlled trials that supported European cancer drug approvals from 2014 to 2016 were assessed to be at high risk of bias based on characteristics of their design, conduct, or analysis; trials that evaluated overall survival were at lower risk of bias than those that evaluated surrogate measures of clinical benefit

Risk of bias judgments differed when using information available from the scientific literature and European regulatory documents separately, which highlights persistent limitations in information available in regulatory documents and published papers

European regulators frequently raised questions about the appropriateness and applicability of the available evidence on new cancer drugs, which were not acknowledged in the scientific literature

Introduction

Regulatory agencies are responsible for evaluating the clinical efficacy and safety of new medicines. In the European Union, the European Medicines Agency (EMA) serves as the gatekeeper to the pharmaceutical market; clinicians can only prescribe a new drug after it receives the EMA's approval.[1] The EMA bases its decisions on a small number of key clinical studies completed and submitted by pharmaceutical manufacturers.[2] Between 2012 and 2016, about half

of new drugs approved by the EMA were associated with a single pivotal study.[3]

Recently, cancer drugs have comprised the single largest category of new drug approvals in Europe. In 2017, more than a quarter (24/92) of EMA approvals were for cancer drugs.[4] There is considerable debate and controversy about the therapeutic and economic value of these drugs.[5-10] Our recent research showed that most new cancer drugs were approved by the EMA without evidence of benefit on overall survival or quality of life.[11] In recent years there has been a substantial shift towards use of surrogate endpoints such as progression free survival.[12] There is growing recognition that the correlation between surrogate endpoints and overall survival is often poor.[13]

About a third of "positive" randomised controlled trials of cancer drugs report treatment effects that are considered to be clinically meaningful according to the European Society of Medical Oncology Magnitude of Clinical Benefit Scale.[5 14] Moreover, there is no association between magnitude of benefit and drug price.[15] Because cancer drugs are responsible for most of the recent increases in pharmaceutical spending across healthcare systems,[16] the evidence base that supports their market entry warrants close scrutiny.

Previous work described the characteristics of pivotal studies supporting new cancer drug approvals in Europe and the United States. Regulators in both settings generally review the same set of clinical studies when approving new drugs.[17] In a large evaluation that focused on US Food and Drug Administration approvals, clinical studies of cancer drugs were less likely to be randomised and double blinded than clinical studies of drugs in other therapeutic areas.[18] In another US study, cancer drugs with orphan (rare disease) designations were less likely to be randomised and double blinded than non-orphan cancer drugs.[19] In Europe, most of the new cancer drug approvals between 2009 and 2013 were supported by at least one randomised controlled trial.[11] However, an increasing proportion of new cancer drugs are approved on the basis of non-randomised, single arm studies.[5]

Randomised controlled trials are widely considered to be the "gold standard" for evaluating the clinical efficacy of new drugs.[20] However, flaws in the design, conduct, analysis, or reporting of randomised controlled trials can produce bias in estimates of treatment effect, potentially jeopardising the validity of their findings. A large body of literature documents these biases, which could be substantial in magnitude.[21-27] For example, in a large meta-epidemiological study of 1973 randomised controlled trials, lack of blinding was associated with an average 22% exaggeration of treatment effects among trials that reported subjectively assessed outcomes.[27] Such non-trivial differences could affect how trial results are interpreted and used in regulatory settings and clinical practice. Therefore, it is imperative to systematically examine the validity of randomised controlled trials that support the approval of new drugs through assessment of the risk of bias in their results.[28]

In this study, we examined the characteristics of randomised controlled trials that supported approval of cancer drugs by the EMA from 2014 to 2016. We focused on three aspects of cancer drug trials. The first aspect was trial design. A key feature of cancer drug trials is whether they are designed to demonstrate a benefit on overall survival or quality of life.[12 29 30] We determined whether recent randomised controlled trials of cancer drugs included overall survival or quality of life outcomes as endpoints. The second aspect was risk of bias. Previous studies primarily focused on crude metrics of trial quality such as blinding of participants and investigators.[18 19 31] Although these aspects are important, they are not an adequate measure of a trial's validity because randomised controlled trials with blinding might still be at high risk of bias (conversely, trials without blinding could produce valid results).[32] We aimed to perform risk of bias assessments that more thoroughly evaluated deficits in the design, conduct, analysis, and reporting of randomised controlled trials.[28] The third aspect was the adequacy, completeness, and consistency of reporting across different sources. Trial reporting has improved substantially over the past few decades.[33] Yet, discrepancies can occur between regulatory documents and scientific publications,[34] and might lead to different interpretations. We investigated such discrepancies.

Methods
Identification of cancer drug approvals
Two researchers (XRS and NH) independently searched the publicly available EMA database of European public assessment reports from 1 January 2014 to 31 December 2016. They used Anatomical Therapeutic Chemical Classification (ATC) codes L01-04 to identify "antineoplastic and immunomodulating" agents for solid tumours and haematological malignancies that received "first marketing authorisations". A third researcher (HN) independently confirmed the sample of cancer drug approvals during this period. We excluded "type 2 variations," which are additional marketing authorisations of already approved drugs in new therapeutic indications.

Our study period ended in 2016, which allowed a minimum of one year for trials to be published in the peer reviewed literature after authorisation. We excluded approvals for the treatment of benign tumours, supportive treatments, and generic products, which was consistent with our previous study.[11]

We noted when a drug received a "conditional marketing authorisation" from the EMA. Conditional marketing authorisations are granted for drugs aimed at treating serious or life threatening conditions with an unmet medical need.[35] Such approvals rely on less comprehensive data than those required for regular marketing authorisations, and pharmaceutical companies are required to conduct additional studies to evaluate the clinical benefit of their products after market entry.[36 37] We also noted if a drug received an "orphan drug" designation, which is granted for the treatment of rare diseases.

Identification of pivotal trials supporting cancer drug approvals

We included the pivotal studies that formed the basis of cancer drug approvals during our study period. Two researchers (XRS and NH) independently identified the pivotal studies, which were defined as those labelled as "main studies" in the European public assessment reports on the EMA website. A third researcher (HN) confirmed the list of pivotal studies. European public assessment reports are summaries of documents compiled by rapporteurs from European member states by using data submitted by pharmaceutical companies. These reports include publicly available information on the characteristics, findings, and EMA's appraisal of pivotal and supportive clinical studies that support marketing authorisation decisions of new products.

Regulatory approval could precede the publication of trial results in the scientific literature. We searched publicly available clinical trial registries in Europe (European Clinical Trials Database: EudraCT) and the US (US National Library of Medicine database of clinical trials: ClinicalTrials.gov) to identify published accounts of pivotal studies in the scientific literature. The latest search date was 15 May 2018. ClinicalTrials.gov routinely searches PubMed and automatically retrieves the peer reviewed publications associated with each record in the registry (study sponsors can also manually enter information on trial publications in ClinicalTrials.gov).[38] Therefore, we primarily relied on ClinicalTrials.gov to identify the publications of pivotal studies supporting EMA approvals during our study period. We cross checked EudraCT to capture any studies which might have been missed in ClinicalTrials.gov. If available, we also identified the protocols and supplementary appendices of pivotal studies. When the protocol was not available, we contacted the corresponding authors and requested access to their study protocol.

Data collection on study design characteristics

We documented the therapeutic indications for which cancer drugs received EMA first marketing authorisations. Pivotal studies were then characterised in terms of their design (randomised *v* non-randomised), study arms (experimental treatment and comparators), and primary and secondary endpoints. We categorised pivotal studies as randomised controlled trials if participants were randomly allocated to different treatment arms. We noted if trial endpoints included overall survival, health related quality of life, progression free survival, disease response rates or response duration. In addition to recording the primary endpoint of each study, we noted whether the secondary endpoints included overall survival or quality of life outcomes. All data were collected independently by two researchers (XRS and NH) and verified by a third (HN).

Risk of bias assessment in randomised trials

We used the revised Cochrane risk of bias assessment tool (RoB 2.0, version 2016, available at www.riskofbias.info) to examine the internal validity of randomised controlled trials.[39] The Cochrane risk of bias tool was initially published in 2008[28] and has been widely used in systematic reviews of randomised controlled trials.[40] The updated version considers five bias domains: (a) bias arising from the randomisation process; (b) bias owing to deviations from intended interventions; (c) bias caused by missing outcome data; (d) bias in measurement of the outcome; and (e) bias in selection of the reported result.[41] Risk of bias judgments were based on answers to a series of signalling questions in each of the five bias domains. We relied on the tool's standard algorithms to map our responses to signalling questions to risk of bias judgments. As recommended in the guidance document, if a trial was judged to be at "high risk of bias" in one domain, we considered it to be at high risk of bias overall.[39] In addition, a trial judged to have "some concerns" in three or more domains was considered to be at high risk of bias overall.[39]

Our assessment focused on the primary endpoint of each pivotal randomised controlled trial. If a trial had clinical and surrogate measures as coprimary endpoints, we relied on the clinical outcome (for example, risk of bias assessment was based on overall survival if a trial included overall survival and progression free survival as coprimary endpoints), unless data were only available for the surrogate measure at the time of EMA approval.

On the first domain, trials were judged to be at low risk of bias if they adopted appropriate methods to generate and conceal the allocation sequence.[42 43] We also examined whether there were imbalances in group sizes, baseline characteristics, or key prognostic factors that suggested a problem with the randomisation process. Previous meta-epidemiological studies have shown that trials with inadequate or unclear sequence generation and allocation concealment have on average 7-10% exaggerated treatment effects compared with trials with adequate methods.[26]

On the second domain, our assessment was based on the effect of assignment to the interventions at baseline (the "intention to treat" effect). Studies were at low risk of bias if there were no deviations from intended interventions. We judged trials to be at high risk of bias if there were clear deviations from the intervention that was intended in the trial protocol; if such deviations were not balanced between the experimental and control groups; and if these deviations likely influenced the outcome. Trials were also judged to be at high risk of bias on this domain if some participants were not analysed in the group to which they were randomised and if there was potential for a substantial impact on study findings. Open label trials were not automatically at high risk of bias owing to deviations from intended interventions; similarly, trials with blinding were not immune to bias by default. In trials that masked participants, carers, and trial personnel, we carefully considered whether blinding could be compromised because of major differences in drug adverse events.[44 45] However, compromised blinding led to high risk of bias judgments for this domain only

if there were deviations from intended interventions that were not balanced between groups and potentially influenced the outcome.

On the third domain, studies were at low risk of bias if outcome data were available for all or nearly all randomised participants, as reported in the CONSORT (Consolidated Standards of Reporting Trials) flow diagram.[46] Unless there was evidence that results were robust to missing outcome data in sensitivity analyses, we considered trials to be at high risk of bias if the proportions of, or reasons for, missing outcome data differed between experimental and control groups (potentially resulting in an imbalance in censoring rates).[47][48] For time to event analyses, trials were judged to be at high risk of bias if, firstly, the proportions of participants who withdrew their consent to take part in the study differed between trial arms, and participants were censored when they withdrew consent; or secondly, participants who discontinued treatment were censored. In these cases, missingness could depend on the true value of the outcome. In some trials, patients who withdrew their consent continued to be followed up and contribute outcome data, unless they specifically withdrew their consent for further tumour assessments. In such cases, we judged the trials to be at low risk of bias.

On the fourth domain, studies were at low risk of bias if outcome assessors were unaware of the intervention received by the study participants. We judged trials to be at high risk of bias if outcome assessors were not masked (or if blinding could be compromised) and outcome assessment could be influenced by knowledge of the intervention received. According to previous meta-epidemiological reviews, lack of blinding of outcome assessors in randomised controlled trials is associated with 36% exaggerated treatment effects for subjective outcomes.[49]

On the fifth domain, studies were at low risk of bias if the results were unlikely to have been selected on the basis of either multiple outcome measurements or multiple analyses of the data.

Two trained researchers (XRS and NH) independently assessed the risk of bias in each pivotal randomised controlled trial. Areas of disagreement were resolved by discussion and consensus in face to face meetings, first between the two researchers and then with other members of the project team (HN and CD). Two researchers reached the same overall risk of bias judgment for 74% of trials when using the published articles and for 85% of trials when using the European regulatory documents. A third researcher (HN) independently reviewed and confirmed the accuracy of all assessments. Difficult cases were also discussed among team members with methodological (JS, JPTH, and JACS) and clinical (BG) expertise.

Reporting adequacy in regulatory documents and scientific literature
We assessed the risk of bias in each randomised controlled trial twice to compare the completeness and consistency of trial reporting in the regulatory documents and scientific literature.[50] Firstly, we relied on the published articles,

and if available, their protocols and supplementary appendices. Secondly, we repeated the assessments by using information available in European public assessment reports alone (without consulting the trial publication, its protocol and supplementary appendix, or clinical trial registry records). There was a minimum "wash out" period of four weeks between assessments. If our risk of bias judgment differed when we used information from regulatory documents versus publications, we noted the reasons for the observed discrepancies. A third version of each risk of bias assessment was derived using the totality of information available for each trial by combining the scientific literature and regulatory documents ("combined information").

Finally, one researcher (HN) documented the additional limitations of the available evidence highlighted in European public assessment reports and whether these were acknowledged in trial publications. These issues were related to the appropriateness and generalisability of the available evidence and focused on trial features that were not included in the five domains of the Cochrane risk of bias tool. For example, regulators often commented on the magnitude of clinical benefit, choice of comparators, and endpoints because these could affect the relevance of trial findings to clinical practice. Similarly, "maturity" of statistical analyses was routinely discussed because early termination of trials could affect the reliability and interpretation of findings, especially when interim analyses are not prespecified.[51][52]

Patient and public involvement
No patients or members of the public were involved in setting the research question or the outcome measures, nor were they involved in developing plans for design or implementation of the study. No patients or members of the public were asked to advise on interpretation or writing up of results. We plan to involve patients and members of the public when disseminating the study results on a publicly available website. We plan to disseminate the findings of this work to patient organisations.

Results
Cancer drug approvals
Figure 1 shows the process that led to identification of our study sample. Of 64 potentially relevant marketing authorisations granted by the EMA between 1 January 2014 and 31 December 2016, 48 were for cancer products. After we excluded 16 generic and supportive care drugs, our sample consisted of 32 cancer drug approvals. A total of five (16%) approvals were indicated for the treatment of multiple myeloma, four (13%) for melanoma, and four (13%) for lung cancer (appendix table 1). During this period, 13 (41%) cancer drugs were approved for orphan indications and five (16%) received conditional marketing authorisations.

Characteristics of pivotal trials
Thirty two cancer drug approvals were supported by 54 pivotal studies. Of these, 41 (76%) were randomised
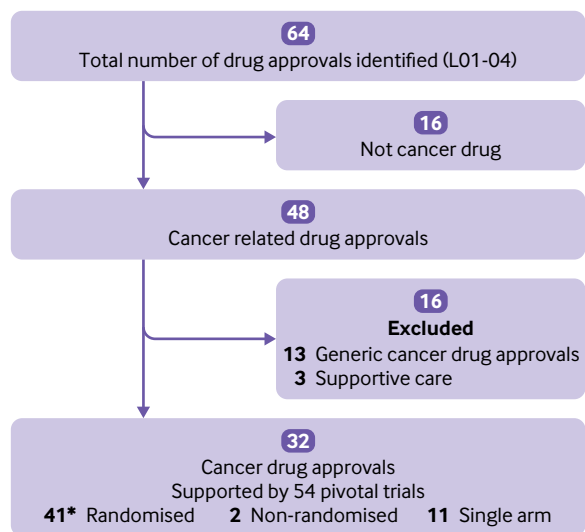
Fig 1 | Identification and selection of cancer drug approvals. *One randomised study corresponds to four separate publications submitted to the European Medicines Agency as bibliographic references. This study was not evaluated further. L01-04=Anatomical Therapeutic Chemical Classification (ATC) codes

controlled trials (two of which randomised patients to different doses of the experimental treatment), 11 (20%) were single arm studies that evaluated the experimental treatment alone without a comparator, and 2 (4%) were non-randomised comparative studies. Only seven cancer drug approvals were supported by two or more randomised controlled trials. Two of 13 cancer drugs (15%) with orphan designation were supported by single arm studies alone compared with 2 of 19 non-orphan drugs (11%). Two of 5 cancer drugs (40%) with conditional marketing authorisations were supported by single arm studies alone, whereas 2 of 27 drugs (7%) with regular approvals had only single arm studies.

We were able to identify published accounts of 39 of 41 randomised controlled trials. The two unpublished randomised controlled trials were for pegaspargase, which was previously approved by national health authorities in several European countries. European public assessment reports were available for all 39 randomised controlled trials. Trial protocols were publicly available for 21 randomised controlled trials and we gained access to two additional protocols.

Only 10 of 39 randomised controlled trials (26%) evaluated overall survival as either a primary or coprimary endpoint (appendix table 2). Progression free survival was the primary endpoint in 21 randomised controlled trials (54%), whereas the remaining trials evaluated disease response, event free survival, or safety endpoints.

Risk of bias
Figure 2 shows the risk of bias in pivotal randomised controlled trials of cancer drugs by using combined information obtained from regulatory documents and the scientific literature. Based on our answers to signalling questions (appendix table 3), we judged two trials to be at high risk of bias that arose from the

randomisation process. Four trials were at high risk of bias owing to deviations from intended interventions. Twenty three randomised controlled trials had some concerns because of deviations from intended interventions, which reflected either lack of blinding or risk of compromised blinding; however, none of these was responsible for a high risk of bias judgment overall. Ten trials were judged to be at high risk of bias owing to missing outcome data and seven because of measurement of the outcome. All randomised controlled trials were at low risk of bias in selection of the reported result.

Taken together, 19 randomised controlled trials (49%) were at high risk of bias overall, 2 (5%) had some concerns, and 18 (46%) were at low risk of bias according to our assessments using the revised Cochrane tool. Of the 19 randomised controlled trials that were at high risk of bias overall, 13 had one domain at high risk of bias, 5 had two domains at high risk of bias, and 1 had three domains with some concerns. Detailed justifications for our judgments are included in appendix table 3.

Fewer randomised controlled trials that evaluated overall survival as the primary or coprimary endpoint were at high risk of bias than those that evaluated surrogate efficacy endpoints (2/10 (20%) v 16/29 (55%), respectively). Of the 16 randomised controlled trials with surrogate endpoints that were at high risk of bias, our judgments were informed by concerns about missing outcome data for six randomised controlled trials; measurement of the outcome for three randomised controlled trials; and a combination of domains for seven randomised controlled trials.

Reporting adequacy
Figure 3 shows the domain specific risk of bias judgments when we considered information reported in the scientific literature (trial publication, protocol, appendix, clinical trial registry record) and European public assessment reports separately. Our judgments differed for at least one domain in 26 out of 39 randomised controlled trials (table 1). Most of these differences did not change the overall risk of bias judgments, however our conclusions changed for eight randomised controlled trials (21%).

Table 1 lists the reasons for observed differences. Overall, the content and consistency of reporting varied between the two sources. For example, the methods adopted in generating and concealing the allocation sequence were more readily available in the scientific literature than in regulatory documents (n=15). In contrast, major protocol deviations were only explicitly reported in regulatory documents, albeit inconsistently (n=3). Although protocols were available for 23/39 randomised controlled trials in our sample, we could not spot major deviations without explicit acknowledgment and discussion of such deviations in the reports. For the remaining bias domains, there was no discernible pattern. While some regulatory documents had more complete reporting in terms of missing outcome data, this information was

Drug name (generic)	Name of study	Primary efficacy endpoint	Randomisation	Deviations from intended interventions	Missing outcome data	Measurement of the outcome	Selection of reported result	Overall risk of bias
Cabozantinib	XL184 -301	PFS	Some concerns	High	High	Low	Low	High
Trametinib	MEK114267	PFS	Low	Some concerns	High	Low	Low	High
Trametinib	BRF113220	Response, PFS	Low	Some concerns	High	High	Low	High
Trametinib	MEK115306	PFS	Low	Low	High	Low	Low	High
Obinutuzumab	BO21004/CLL11	PFS	Low	Some concerns	Low	Low	Low	Low
Idelalisib	GS -US-312-0116	PFS	Low	Low	High	Low	Low	High
Ibrutinib	Study 1112	PFS	Low	Some concerns	High	Low	Low	High
Nintedanib	LUME Lung 1— 1199.13	PFS	Low	Low	High	Low	Low	High
Olaparib	D0810C00019	PFS	Low	Low	Low	Low	Low	Low
Ramucirumab	I4T-IE-JVBD (REGARD)	OS	Low	Low	Low	Low	Low	Low
Ramucirumab	I4T-IE-JVBE (RAINBOW)	OS	Low	Some concerns	Low	Low	Low	Low
Lenvatinib	E7080-G000-303 (SELECT)	PFS	Low	Some concerns	Low	Low	Low	Low
Nivolumab	CA209-066	OS	Low	Low	Low	Low	Low	Low
Nivolumab	CA209-037	Response, OS	Low	Some concerns	High	Low	Low	Low
Pembrolizumab	Study P002	PFS, OS	Low	Some concerns	Low	Low	Low	Low
Pembrolizumab	Study P006	PFS, OS	Low	Low	Low	Low	Low	Low
Sonidegib	CLDE225A2201 (BOLT)	Response	Low	Low	High	Low	Low	High
Panobinostat	CLBH589D2308 (Panorama I)	PFS	Low	Some concerns	Low	Low	Low	Low
Carfilzomib	PX-171-009 (ASPIRE Study)	PFS	Low	High	Low	Low	Low	High
Cobimetinib	GO28141 (coBRIM)	PFS	Low	Low	Low	Low	Low	Low
Talimogene laherparepvec	Study 005/05	Response	Low	Some concerns	Low	Some concerns	Low	Some concerns
Pegaspargase	CCG-1962	No efficacy	Some concerns	Some concerns	Low	Low	Low	Some concerns
Pegaspargase	DFCI-87-001	EFS	Low	Some concerns	High	High	Low	High
Pegaspargase	DFCI-91-01	EFS	Low	Some concerns	Low	High	Low	High
Pegaspargase	DFCI-05-001	DFS, OS, HRQoL	Low	Low	Low	High	Low	High
Pegaspargase	AALL07P4	No efficacy	Some concerns	Low	Some concerns	Low	Low	Some concerns
Pegaspargase	ASP-304	Response	High	Some concerns	Low	High	Low	High
Necitumumab	I4X-IE-JFCC (SQUIRE)	OS	Low	Some concerns	Low	Low	Low	Low
Trifluridine/tipiracil	TPU-TAS-102-301 (RECOURSE)	OS	Low	Low	Low	Low	Low	Low
Elotuzumab	CA204-004	PFS, Response	Low	Some concerns	Low	Low	Low	Low
Elotuzumab	CA204-009	PFS	Low	Low	Low	High	Low	High
Daratumumab	MMY2002	Response	High	High	Low	Low	Low	Low
Lenvatinib	E7080-G000-205	PFS	Low	Low	Low	Low	Low	Low
Cabozantinib	XL184-308	PFS	Low	Low	Low	Low	Low	Low
Irinotecan hydrochloride trihydrate	NAPOLI-1	OS	Low	Some concerns	Low	Low	Low	Low
Olaratumab	I5B-IE-JGDG (JGDG)	PFS	Low	Some concerns	Low	Low	Low	Low
Palbociclib	1023 (PALOMA-3)	PFS	Low	Some concerns	Low	High	Low	High
Palbociclib	1008 (PALOMA-2)	PFS	Low	High	Low	Low	Low	High
Ixazomib	C16010	PFS	Low	Low	Low	Low	Low	Low

● High ● Some concerns ● Low

Fig 2 | Risk of bias assessment for the pivotal randomised controlled trials supporting European Medicines Agency cancer drug approvals using information available in the scientific literature (trial publications, protocols, supplementary materials, and clinical trial registry records) and regulatory documents. Risk of bias assessments were based on the primary efficacy endpoints. DFS=disease free survival; EFS=event free survival; HRQoL=health related quality of life; OS=overall survival; PFS=progression free survival
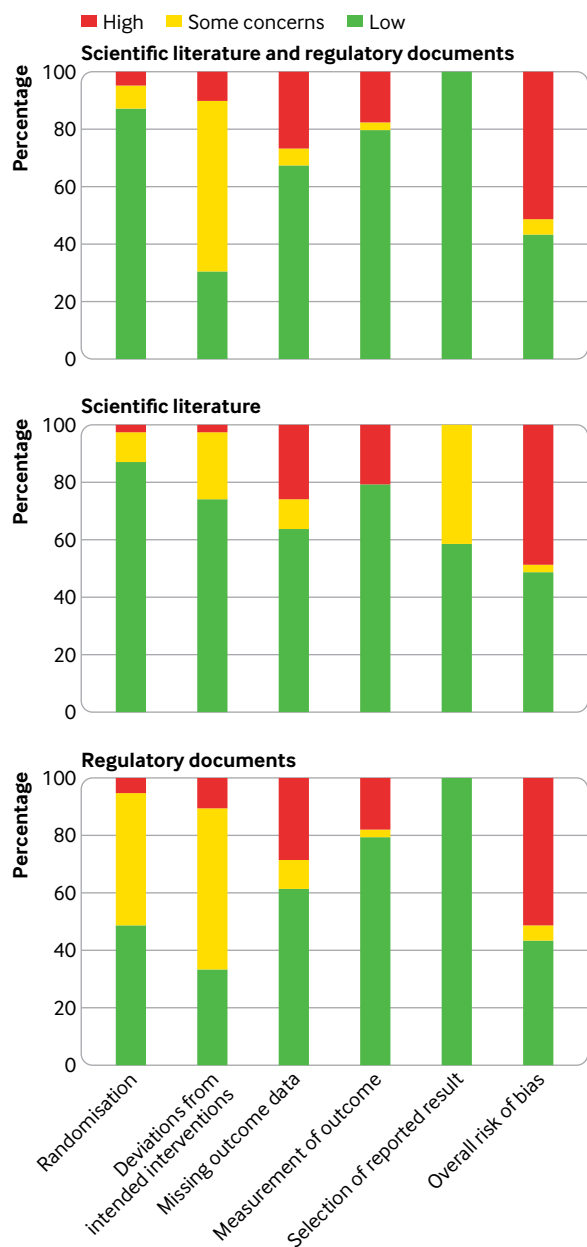
Fig 3 | Risk of bias assessments using combined information from the scientific literature and regulatory documents, only information available in the scientific literature, and only information available in regulatory documents. Risk of bias assessments were based on the primary efficacy endpoints

more routinely and comprehensively reported in the scientific literature in other cases.

Regulatory reviewers identified additional limitations in the evidence base beyond the domains captured in the risk of bias assessments for 10 drugs (32%). These limitations focused on the appropriateness and generalisability of the available data and included choice of comparators, study endpoints, interim analyses, or a combination of factors (box 1). In five cases, the committee questioned either the consistency or magnitude of the observed clinical benefit. The regulatory reviewers raised concerns that were substantial enough to warrant a divergent committee opinion in four cases, all of which were judged to be at

low risk of bias according to our risk of bias assessment. None of these regulatory concerns were fully disclosed and discussed as limitations or uncertainties in the scientific literature (appendix box 1).

Results summary

Table 2 summarises our findings at the cancer drug level. Of 32 new cancer drugs approved by the EMA from 2014 to 2016, 27 entered the European market with at least one randomised trial. Of the cancer drugs with randomised controlled trials, only seven were evaluated in trials powered to measure overall survival as a primary or coprimary endpoint. Half (n=16) of cancer drugs had at least one randomised controlled trial at low risk of bias. European regulators identified other concerns for 7 of the 16 drugs that had at least one randomised controlled trial at low risk of bias.

Figure 4 summarises our findings according to approval characteristics. Of 13 cancer drugs approved in orphan conditions, 4 (31%) had at least one randomised controlled trial at low risk of bias and without major regulatory concerns. The corresponding number was 5 among the subset of 19 drugs (26%) approved in non-orphan conditions. A lower proportion of cancer drugs with conditional marketing authorisations had at least one randomised controlled trial at low risk of bias and without major regulatory concerns compared with drugs with regular EMA approvals (1/5 (20%) v 8/27 (30%), respectively).

Discussion

Summary of findings

In this study, we evaluated the evidence base underpinning the EMA's recent cancer drug approvals. Between 2014 and 2016, a quarter of pivotal studies supporting cancer drug approvals were not randomised designs. Of the 39 randomised controlled trials that formed the basis of new cancer drug approvals, almost three quarters did not measure overall survival or quality of life outcomes as primary endpoints. Using the revised Cochrane tool, we judged 49% of randomised controlled trials to be at high risk of bias. Our judgments changed in either direction for a fifth of randomised controlled trials when we relied on information reported in regulatory documents and scientific publications separately. Regulators identified additional deficits beyond the domains captured in risk of bias assessments for several trials, which were not disclosed as limitations in scientific publications.

The three key findings of this study warrant further discussion. Firstly, our evaluation characterises the design features of contemporary cancer drug trials. Although randomised controlled trials accounted for about 90% of pivotal studies from 2009 to 2013,[11] such designs accounted for 75% of studies from 2014 to 2016. A growing proportion of recent cancer drug approvals were based on single arm studies, which are more likely to receive conditional marketing authorisations that target indications with unmet medical need.[53] Even when trials had comparators, their appropriateness was at times questionable.

Table 1 | Reasons for differences in risk of bias judgments according to scientific literature and regulatory documents (only randomised controlled trials with different risk of bias judgments are listed)

Drug name (generic)	Name of study	Observed discrepancy	Reason for discrepancy
Cabozantinib	XL184-301	Deviation from intended interventions: some concerns in published reports and high risk of bias in EPAR	EPAR provided additional detail on major protocol deviations
Trametinib	BRF113220	Randomisation: low risk of bias in published reports and high risk of bias in EPAR. Missing outcome data: some concerns in published reports and high risk of bias in EPAR	EPAR lacked information on randomisation methods. EPAR provided additional detail on (non-event related) censoring of participants
Trametinib	MEK115306	Randomisation: low risk of bias in published reports and high risk of bias in EPAR. Missing outcome data: low risk of bias in published reports and high risk of bias in EPAR	EPAR lacked information on randomisation methods. EPAR provided additional detail on the number of participants with censoring
Idelalisib	GS-US-312-0116	Missing outcome data: low risk of bias in published reports and high risk of bias in EPAR	Numbers of participants who withdrew their consent do not match between the participant disposition reported in EPAR and supplementary appendix of publication
Ibrutinib	Study 1112	Randomisation: low risk of bias in published reports and some concerns in EPAR	EPAR lacked information on randomisation methods
Nintedanib	LUME Lung 1 – 1199.13	Randomisation: low risk of bias in published reports and some concerns in EPAR	EPAR lacked information on randomisation methods
Olaparib	D0810C00019	Randomisation: low risk of bias in published reports and some concerns in EPAR	EPAR lacked information on randomisation methods
Ramucirumab	I4T-IE-JVBD (REGARD)	Randomisation: low risk of bias in published reports and some concerns in EPAR	EPAR lacked information on randomisation methods
Sonidegib	CLDE225A2201 (BOLT)	Randomisation: low risk of bias in published reports and some concerns in EPAR	EPAR lacked information on randomisation methods
Panobinostat	CLBH589D2308 (Panorama I)	Missing outcome data: low risk of bias in published reports and high risk of bias in EPAR	Supplementary appendix of published report included results from sensitivity analyses that tested different censoring rules. EPAR did not report these results
Carfilzomib	PX-171-009 (ASPIRE Study)	Deviation from intervention intended interventions: some concerns in published reports and high risk of bias in EPAR	EPAR reported the proportion of participants with major protocol deviations, which differed across trial arms
Talimogene laherparepvec	Study 005/05	Missing outcome data: high risk of bias in published reports and low risk of bias in EPAR. Measurement of the outcome: low risk of bias in published reports and some concerns in EPAR	EPAR reported findings of sensitivity analyses that tested different censoring rules (not available in publication). EPAR provided additional information on outcome assessment
Pegaspargase	CCG-1962	Missing outcome data: some concerns in published report and low risk of bias in EPAR	EPAR provided more information about study and treatment discontinuations
Pegaspargase	DFCI-87-001	Randomisation: low risk of bias in publication and some concerns in EPAR. Missing outcome data: low risk of bias in published report and high risk of bias in EPAR	EPAR lacked information on randomisation methods. Number of randomised and assessed participants did not match between publication and EPAR
Pegaspargase	DFCI-91-01	Randomisation: low risk of bias in publication and some concerns in EPAR	EPAR lacked information on randomisation methods
Pegaspargase	DFCI-05-001	Randomisation: low risk of bias in published report and some concerns in EPAR. Missing outcome data: low risk of bias in publication and some concerns in EPAR	Published report provided additional information about availability of outcome data
Pegaspargase	AALL07P4	Missing outcome data: high risk of bias in published report and some concerns in EPAR	Published report provided additional information about availability of outcome data
Pegaspargase	ASP-304	Missing outcome data: some concerns in published report and low risk of bias EPAR	EPAR provided more detailed information about outcome data availability
Elotuzumab	CA204-009	Randomisation: some concerns in published reports and low risk of bias in EPAR. Deviation from intended intervention: some concerns in publication and low risk of bias in EPAR. Missing outcome data: low risk of bias in published reports and high risk of bias in EPAR	EPAR provided slightly more information on randomisation methods. EPAR reported that major protocol deviations were balanced between trial arms. Findings of sensitivity analyses that tested different censoring rules were available only in published reports
Daratumumab	MMY2002	Randomisation: low risk of bias in published reports and high risk of bias in EPAR	EPAR provided more detail on different parts of the trial, which were not randomised
Lenvatinib	E7080-G000-205	Deviation from intended intervention: some concerns in published report and low risk of bias in EPAR. Missing outcome data: high risk of bias in published report and low risk of bias in EPAR. Measurement of the outcome: high risk of bias in published report and low risk of bias in EPAR	EPAR reported that protocol deviations were balanced across trial arms. EPAR also reported the findings of sensitivity analyses testing different censoring rules. Results of outcome assessment by an independent panel were reported in EPAR
Cabozantinib	XL184-308	Randomisation: low risk of bias in published report and some concerns in EPAR. Deviation from intended intervention: some concerns in published report and low risk of bias in EPAR	EPAR lacked information on randomisation methods. EPAR provided information on protocol deviations, which appeared balanced across trial arms
Olaratumab	I5B-IE-JGDG (JGDG)	Missing outcome data: high risk of bias in published report and low risk of bias in EPAR	EPAR reported findings from sensitivity analyses that tested different censoring rules
Palbociclib	1023 (PALOMA-3)	Randomisation: low risk of bias in published report and some concerns in EPAR	EPAR lacked information on randomisation methods
Palbociclib	1008 (PALOMA-2)	Randomisation: low risk of bias in published report and some concerns in EPAR. Deviation from intended intervention: some concerns in published report and high risk of bias in EPAR	EPAR lacked information on randomisation methods. EPAR reported major protocol deviations, which differed across trial arms
Ixazomib	C16010	Randomisation: low risk of bias in published report and some concerns in EPAR	EPAR lacked information on randomisation methods

EPAR=European public assessment report.

Box 1: Major concerns raised in European regulatory documents (appendix box 1 includes a more detailed narrative overview of these cases)

Idelalisib: early termination of trial
- The pivotal study was terminated early owing to efficacy. According to the EPAR, "the magnitude of the treatment effect is therefore not well defined and further follow-up is needed."
- This need was acknowledged in the published article for safety evaluation but not to confirm clinical benefit.

Olaparib: non-preferred outcome; retrospective subgroup analysis
- Regulators had previously advised the manufacturer to evaluate overall survival as the trial's primary endpoint; however, this recommendation was not followed.
- The target patient population for which approval was sought was retrospectively defined. In addition, "the SAG was uncertain about the true effect of olaparib in this [target] patient population due to the shortcomings of the pivotal study being a small phase II randomised study with a large percentage of censored observations for PFS analysis, and in view of the absence of improvement in OS."
- These limitations were not acknowledged in the published article. However, the published article reported that: "Our data cannot address differences that might exist between patients with BRCA germline mutations and those with a BRCAness phenotype."

Ramucirumab: inconsistent effects across settings and demographic characteristics; divergent committee opinion
- According to the EPAR, there was a differential OS outcome across regions in the REGARD study. "Inconsistency was also observed regarding gender (ramucirumab effective in men, but potentially detrimental in women)."
- In addition, some CHMP members "considered that the effect associated with ramucirumab as single agent was too marginal and possibly even inferior to single agent chemotherapy that are used in this setting."
- In contrast, the published article for the REGARD study stated that: "the survival benefit with ramucirumab was consistent across almost all subgroups. Although the effect on OS was attenuated in women, the PFS estimate in women favoured ramucirumab."

Nivolumab: unplanned early termination of trial due to observed efficacy; non-preferred outcome; potential imbalances in baseline characteristics
- According to the EPAR, "the CHMP was concerned that … carrying out an unplanned interim analysis … was questionable and would introduce uncertainties such as the potential for informative bias." While the published article specified that this analysis was "unplanned," it did not comment on the potential implications of this decision.
- The CHMP highlighted that use of ORR as a primary endpoint in study CA209-037 was not recommended and "OS should have been favoured as primary endpoint …"
- In addition, "the CHMP had concerns over the shape of the OS curves … The applicant provided a discussion on the possible confounding factors, mainly that there were baseline imbalances between the nivolumab arm and chemotherapy arm."
- Although the published article reported that "baseline characteristics were similar in the nivolumab and [investigators' choice chemotherapy] study groups, with the exception of history of brain metastases and high lactase dehydrogenase, which were higher in the nivolumab than the [investigators' choice chemotherapy group]," findings on OS were not reported.

Panobinostat: questionable clinical benefit
- According to the EPAR, the CHMP concluded that "the benefit in PFS has not been translated into a similar relative benefit in OS." A SAG meeting was convened to discuss whether the clinical benefit of panobinostat was "sufficient to justify exposing these patients to the severe adverse event profile of the drug."
- Some members of the SAG concluded that "the clinical benefit cannot be considered established." Such concerns were not disclosed or discussed in the published article.
- While the EPAR showed that "the results from the EORTC QLQ-C30 captured a consistently negative effect by the experimental regimen compared to the control arm …," these results were not reported in the primary publication of the trial.

Cobimetinib: appropriateness of comparator arms
- According to the EPAR, "the CHMP expressed some concern over the lack of a cobimetinib only treatment arm."
- Appropriateness of the comparator arm was not discussed in the published article.

Talimogene: appropriateness of comparator arms; questionable clinical benefit; questionable study endpoints; divergent committee opinion
- According to the EPAR, there were several concerns with the pivotal study supporting this approval.
- The CAT expressed concerns over the comparator; questioned "the validity of using DRR as the primary endpoint for the pivotal trials as opposed to using other more robust endpoints such as PFS or OS"; and questioned the clinical relevance of the magnitude of treatment effect on DRR.
- In addition, "although there appeared to be an effect on OS in the subgroup of patients with Stage IIIB-IVM1a disease, OS was a secondary endpoint and the effect was based on exploratory subgroup analyses, after the analysis in the full analysis set was not statistically significant, and without a pre-specified strategy for multiplicity adjustment."
- These concerns led to a divergent opinion on the approval decision.

Box 1: Continued

- Despite the statement in the EPAR that "it cannot be concluded that an effect on OS has been established for talimogene," the published article reported that it was "well tolerated and resulted in a higher DRR (P<.001) and longer median OS (P=.051) ..." The discussion section added, "combined with the limited toxicity observed, these are clinically important results."

Necitumumab: questionable clinical benefit; divergent committee opinion

- The magnitude of clinical benefit was questioned by the CHMP, which led to a divergent opinion on the approval decision.
- In contrast, the study publication reported that "the addition of a targeted agent to a platinum-based doublet improves survival. These efficacy data and the acceptable safety profile of necitumumab suggest a favourable benefit-to-risk ratio for this combination treatment."

Daratumumab: appropriateness of comparator arms

- According to the EPAR, "the design of the study with no comparative arm is of concern ..."
- This was not highlighted in the published account of the trial. Instead, the publication reported: "although this study did not have a control arm, patients with the degree of treatment refractoriness in our study historically have poor outcomes."

Ixazomib: problematic interim analyses; questionable clinical benefit; divergent committee opinion

- The EMA initially refused granting a marketing authorisation to ixazomib, but this decision was subsequently overturned. Ultimately, ixazomib received a conditional marketing authorisation.
- There were several concerns regarding the single pivotal trial submitted by the manufacturer. Notably, there was a "worsening of results between the initially submitted analysis and the updated analysis."
- The CHMP therefore concluded that ixazomib "was not approvable based on the efficacy grounds ..."
- In response to the manufacturer's request for reexamination, a SAG meeting was convened. The SAG "considered that on the basis of the primary PFS analysis, which was conducted according to the pre-specified statistical considerations, the trial met its objective of showing a statistically and clinically significant improvement in PFS." Although the CHMP still maintained that "the total available evidence on efficacy is not as comprehensive as normally would be required," a conditional marketing authorisation was ultimately granted.
- There was a divergent opinion on the approval decision (signed by nine members of the CHMP).
- The published article only reported the first set of interim results, which were initially submitted to the EMA and stated that "in accordance with the statistical analysis plan in the protocol and the principle of group sequential design, this was the final statistical analysis of progression free survival" despite providing more mature data on progression free survival data to the European regulators.

BRCA=breast cancer gene; CAT=Committee for Advanced Therapies; CHMP=Committee for Medicinal Products for Human Use; DRR=durable disease response; EMA=European Medicines Agency; EORTC= European Organisation for Research and Treatment of Cancer; EPAR=European public assessment report; ORR=overall response rate; OS=overall survival; PFS=progression free survival; SAG=Scientific Advisory Group.

We found two randomised controlled trials in which participants were randomised to receive different doses of the same experimental treatment (without a control). In other cases, the comparator either precluded isolation of the effect of the experimental treatment or did not adequately reflect standard of care; these trials were subsequently criticised by the EMA. In terms of study endpoints, only a quarter of randomised controlled trials were powered to evaluate overall survival as the primary outcome. According to the recent EMA guidelines on the evaluation of anticancer treatments,[54] "convincingly demonstrated favourable effects on overall survival are from both a clinical and methodological perspective the most persuasive outcome of a clinical trial." Yet, most cancer drugs were approved on the basis of other endpoints, such as progression free survival and disease response. Recent systematic reviews showed that progression free survival and disease response do not consistently translate to survival gains or quality of life benefits.[13 55-58] Cancer drugs that appear effective on these surrogate measures could even turn out to be harmful.[59]

Secondly, the evidence base underpinning EMA approvals of new cancer drugs has methodological weaknesses. In this study, the primary domains responsible for the high risk of bias judgments were missing outcome data and measurement of the outcome (see box 2 for illustrative examples). In several trials, the proportions and reasons for missing outcome data differed, which probably resulted in unbalanced censoring[60] and potentially favoured the experimental drug.[61] Considerable differences in the toxicity profiles of drugs was another common issue. When such differences were substantial, we concluded that blinding of participants, carers, or investigators could be compromised. As recognised by EMA scientists, "the real effectiveness of the blinding for cancer drugs can always be questioned."[62] We judged trials to be at high risk of bias if a subjective primary outcome (such as progression free survival) was assessed by local investigators who we concluded might no longer be blinded to treatment allocation. This judgment was also supported by the recent US Food and Drug Administration guidelines that recommend independent verification of tumour assessment endpoints when the adverse event profiles of comparator treatments could substantially unblind the trial in practice.[63]

Thirdly, the regulatory documents and scientific publications had limitations in their reporting. Some

Table 2 | Overview of findings at the drug level. For each cancer drug approval from 2014 to 2016, the table shows whether there was at least one randomised controlled trial supporting the EMA's approval decision; whether there was at least one randomised controlled trial evaluating overall survival as a primary or coprimary endpoint; whether there was at least one randomised controlled trial at low risk of bias; whether EMA scientists and committee members raised additional concerns about the appropriateness of the available evidence according to factors that were not captured in risk of bias assessments; and whether EMA committee members issued a divergent opinion on the approval decision based on those concerns

Drug name (generic)	Approval characteristics	Has at least one RCT supporting EMA approval	Has at least one RCT evaluating OS for EMA approval	Has at least one RCT at low risk of bias	Other regulatory concerns raised prior to approval	Divergent regulatory opinion on the approval
Cabozantinib	Orphan; conditional	Yes	No	No	No	No
Trametinib	—	Yes	No	No	No	No
Obinutuzumab	Orphan	Yes	No	Yes	No	No
Idelalisib	—	Yes	No	No	Yes	No
Ibrutinib	Orphan	Yes	No	No	No	No
Nintedanib	—	Yes	No	No	No	No
Olaparib	Orphan	Yes	No	Yes	Yes	No
Ramucirumab	—	Yes	Yes	Yes	Yes	Yes
Ceritinib	—	No	NA	NA	NA	NA
Lenvatinib	Orphan	Yes	No	Yes	No	No
Nivolumab	—	Yes	Yes	Yes	Yes	No
Pembrolizumab	—	Yes	Yes	Yes	No	No
Sonidegib	—	Yes	No	No	No	No
Panobinostat	Orphan	Yes	No	Yes	Yes	No
Carfilzomib	Orphan	Yes	No	No	No	No
Cobimetinib	—	Yes	No	Yes	Yes	No
Blinatumomab	Orphan; conditional	No	NA	NA	NA	NA
Talimogene	—	Yes	No	No	Yes	Yes
Asparaginase	—	No	NA	NA	NA	NA
Pegaspargase	—	Yes	Yes	No	No	No
Osimertinib	—	No	NA	NA	NA	NA
Necitumumab	—	Yes	Yes	Yes	Yes	Yes
Trifluridine/tipiracil	—	Yes	Yes	Yes	No	No
Elotuzumab	—	Yes	No	Yes	No	No
Daratumumab	Orphan	Yes	No	No	Yes	No
Lenvatinib	—	Yes	No	Yes	No	No
Cabozantinib	—	Yes	No	Yes	No	No
Irinotecan	Orphan	Yes	Yes	Yes	No	No
Olaratumab	Orphan; conditional	Yes	No	Yes	No	No
Palbociclib	—	Yes	No	No	No	No
Ixazomib	Orphan; conditional	Yes	No	Yes	Yes	Yes
Venetoclax	Orphan; conditional	No	NA	NA	NA	NA

EMA=European Medicines Agency; NA=not applicable; OS=overall survival; RCT=randomised controlled trial.
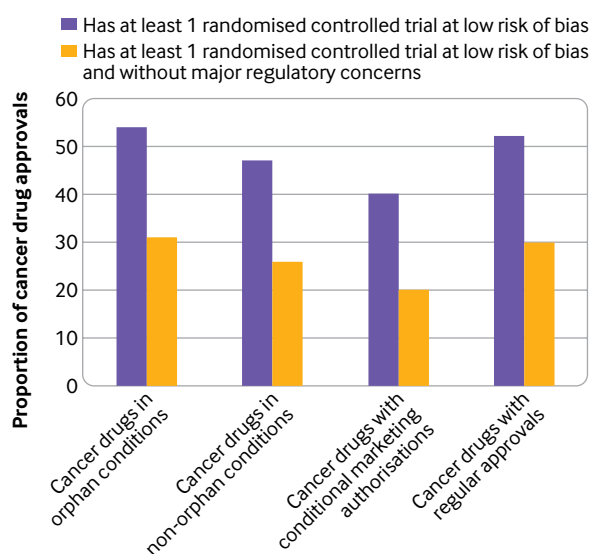
Fig 4 | Overview of findings at the drug level according to approval characteristics (orphan v non-orphan conditions; conditional marketing authorisations v regular approvals). The figure shows whether there was at least one randomised controlled trial at low risk of bias; and whether there was at least one randomised controlled trial at low risk of bias and without major regulatory concerns before approval

of these limitations followed a discernible pattern. For example, key design elements of randomised controlled trials such as sequence generation and allocation concealment were consistently reported in trial publications, their protocols, or appendices. In contrast, regulatory documents seldom specified randomisation methods; instead, regulators often discussed potential imbalances in baseline characteristics to gauge the success of randomisation.[64] Other discrepancies among regulatory documents and publications were less predictable. While both sources typically showed the flow of participants, strategies to deal with missing outcome data and the sensitivity of findings to censoring rules and assumptions were only haphazardly reported. Moreover, neither source consistently reported major deviations from intended interventions.

Comparison to other studies in the literature

Previous studies have documented the shift in cancer trials away from evaluating overall survival in the 1970s to measuring surrogates of clinical benefit in more recent decades.[29 65 66] Our findings confirm that this trend has continued for trials that informed regulatory

Box 2: Illustrative examples of trials judged to be at high risk of bias because of missing outcome data and measurement of the outcome

Elotuzumab

We judged one of the pivotal trials of elotuzumab to be at high risk of bias. In trial CA204-009, which was open label, outcome assessors were aware of the intervention received by study participants, and assessment of the progression free survival outcome could have been influenced by knowledge of the intervention received. Because there was no blinded central assessment of outcomes, we concluded that this trial was at high risk of bias in measurement of the outcome. This potential limitation was also acknowledged in the trial publication.

Nivolumab

We judged one of the pivotal trials of nivolumab to be at high risk of bias because of missing outcome data. In trial CA209-037, outcome data were potentially missing for a considerable proportion of the population. In the investigator's choice chemotherapy arm of the trial, 22/133 (16.5%) patients withdrew their consent, which meant withdrawing consent from the full protocol, including study treatment, study procedures, and survival follow-up. Proportions of missing outcome data and reasons for missing outcome data differed across intervention groups: 16.5% v <1% of patients withdrew their consent in the investigator's choice chemotherapy and nivolumab arms of the trial, respectively. Missingness in the outcome could be related to both the intervention group and the true value of the outcome. Also, there were no sensitivity analyses conducted to test the robustness of study results to different assumptions about missing outcome data.

Trametinib

We judged the pivotal trial of trametinib to be at high risk of bias. In trial BRF113220, there was potential evidence of unbalanced censoring. According to the European public assessment report, there were relatively large proportions of censored participants in both trial arms. The censoring method included censoring for extended loss to follow-up, new anticancer therapy, and excluding symptomatic progression. While 31% of participants who received dabrafenib 150 mg were censored, 11% were censored among participants receiving trametinib. In the absence of evidence that results were robust to the presence of potentially missing outcome data (we were unable to find the results of "eight sensitivity analyses planned to investigate the robustness of progression-free survival against these censoring rules"), we concluded that this trial was at high risk of bias due to missing outcome data. In addition, local investigators in trial BRF113220 were unblinded and therefore aware of the intervention received by study participants. Because the assessment of the progression free survival outcome could be influenced by knowledge of the intervention received, the trial was also at high risk of bias owing to measurement of the outcome. The authors reported local results as their main analysis, although the results from the blinded review committee were also available in the main body of the publication. Results obtained from blinded assessment were less pronounced (hazard ratio for progression free survival was 0.39, 95% confidence interval 0.25 to 0.62, according to local assessment, compared with 0.55, 0.33 to 0.93).

decisions from 2014 to 2016. Similar to recent evaluations, we found low risk of bias arising from the randomisation process and selection of the reported result.[33][67] Our concerns about missing outcome data were supported by an earlier study which showed that about one third of breast cancer trials had differential rates and reasons for censoring.[61] Finally, our findings concur with those from previous studies which showed that incorporating data from additional documents often improves risk of bias assessments.[68-71]

Implications for practice and policy

Our findings highlight the need to improve the design, conduct, analysis, and reporting of cancer drug trials.[72] Regulatory agencies and their evidence requirements shape the design features of pivotal trials.[73-75]

Therefore, regulatory action is needed to ensure that pharmaceutical manufacturers routinely evaluate their products in randomised trials that collect data on meaningful outcomes. In the absence of such data, it remains difficult to know whether new cancer drugs meet the needs of patients, clinicians, and healthcare systems.

While some of the methodological problems identified in our study were avoidable (for example, by ensuring adequate sequence generation and allocation concealment), others could be less straightforward to address in complex cancer trials. For example, ensuring outcome data availability when participants withdraw their consent might not be possible. The proportions of participants who withdrew their consent frequently differed among trial arms, which likely reflected meaningful differences in toxicity profiles of comparator drugs. In such instances when missingness could depend on the true value of the outcome, it is essential to evaluate the sensitivity of trial findings to different assumptions about missing outcome data.[76] However, some trials censored participants when they changed from their assigned treatment. This analysis strategy is not appropriate when estimating intention to treat effects and could lead to bias because of missing outcome data.

Addressing other methodological problems in cancer trials could be feasible, but they come at a cost. Strategies to prevent unblinding might add complexity to trial designs.[23][77] For example, methods to avoid unblinding in randomised controlled trials include centralised dosage modification of treatments and centralised assessment of clinical side effects.[77] Similarly, independent clinicians could perform blinded central evaluation of tumour assessment endpoints, but this might have major cost implications.[78] In a recent review of randomised controlled trials in solid tumours, there was no systematic bias between the findings from blinded independent central review and local assessment, but there were statistical inconsistencies between the two sets of results in almost a quarter of trials.[79] Moreover, previous meta-epidemiological reviews across different therapeutic areas have found that studies with non-blinded assessors of subjective outcomes generate biased findings.[49][80] This research strengthens the argument in favour of implementing blinded centralised assessments of tumour endpoints despite the associated costs.

An important design consideration in cancer trials is the choice of primary endpoint. Surrogate measures of clinical benefit (eg, progression free survival and disease response) have important feasibility advantages because they can be assessed earlier and with smaller sample sizes (and therefore fewer resources) compared with overall survival. A recent study found that cancer drugs approved on the basis of surrogate measures had on average an 11 month shorter development duration compared with drugs approved on the basis of overall survival.[81] However, the feasibility advantages of using surrogate measures

Box 3: Recommendations to improve the design, conduct, analysis, and reporting of pivotal trials of new cancer drugs

- Because overall survival would be largely immune to several sources of potential bias, regulators should require overall survival to be the primary endpoint of pivotal trials.
- Other desirable trial endpoints include quality of life and measures with established surrogacy.
- The magnitude of benefit associated with new cancer therapies should be carefully considered in trial design.
- When subjectively assessed outcomes are used as primary endpoints, trial sponsors should implement blinded independent central review of tumour assessments.
- If the trial is blinded, trial sponsors and investigators should adopt strategies to avoid unblinding of investigators, for example centralised dosage modification of treatments; this is especially important when outcome assessment is not blinded.
- Trial sponsors and investigators should report the risk of unblinding in trials in which investigators are not aware of treatment allocation; this is especially important when outcome assessment is not blinded.
- Trial sponsors and investigators should conduct sensitivity analyses to evaluate the robustness of trial results to missing outcome data. Regulators and journal editors should require consistent reporting of the findings of these sensitivity analyses.
- Regulators and journal editors should require consistent reporting of any major deviations from intended interventions that arose from the experimental context; whether deviations are balanced between the groups; and whether deviations could affect the outcome.

should be weighed against their several disadvantages. Firstly, patients might misinterpret such endpoints and overestimate the magnitude of benefit associated with new cancer drugs.[82 83] Secondly, the strength of the correlation between surrogate and clinical outcomes in cancer trials is unclear.[13] Over the past decade, several drugs (eg, bevacizumab in metastatic breast cancer)[84] approved on the basis of surrogate measures failed to demonstrate overall survival gains in subsequent trials. In the recent BELLINI trial, patients who received venetoclax had shorter survival than those who received a control treatment (even though venetoclax appeared more effective than the control on the basis of progression free survival and response rate).[59] Thirdly, and as recommended by regulators, unblinded trials (or trials at risk of unblinding) with surrogate endpoints might require additional (costly) safeguards such as independent blinded endpoint review to minimise risk of bias.[63] Finally, and perhaps most importantly, evidence of overall survival benefit might never emerge for cancer drugs approved on the basis of surrogate measures alone.[85] In an earlier study, we found that data on overall survival did not emerge in the postmarketing period for more than 90% of indications for which there was no evidence of such a benefit at the time of marketing authorisation.[11]

Taken together, these findings support more widespread use of overall survival as the primary endpoint in pivotal trials of new cancer drugs. Randomised controlled trials with overall survival endpoints were less likely to be at high risk of bias in our sample; this finding is consistent with an earlier assessment that showed that 66% of randomised controlled trials evaluating overall survival were at low risk of bias[68] (the corresponding figure in our study was 80%). Overall survival would be largely immune

to the risk of bias attributable to potential unblinding of outcome assessors and missing outcome data.[22 26 27]

There is also an opportunity to further improve the reporting standards of regulatory documents and scientific publications. Publication of the CONSORT statement in 1996 (and its update in 2010)[46] has led to major improvements in randomised controlled trial reporting in the scientific literature.[86] Also, the 2015 revision[87] of the European public assessment report template addressed some of the previous criticisms.[88] Currently, publications and regulatory documents make it difficult to distinguish between trial deficits that can be avoided and those that are more difficult to address. When methodological shortcomings are inevitable (eg, missing outcome data owing to withdrawal of participant consent), more transparent reporting is warranted. In addition, key information required to perform risk of bias assessments is still inconsistently reported in regulatory documents, trial protocols, publications, supplementary appendices, and clinical trials registries. For example, neither journal articles nor regulatory documents discuss the possibility that trial investigators could be unblinded when the adverse event profiles of comparator treatments are substantially different. Similarly, these sources do not consistently report the occurrence of protocol deviations that arose from the experimental context; whether deviations are balanced between the groups; and whether deviations could affect the outcome. Journal editors and European regulators can take further action to facilitate more complete and consistent reporting of pivotal studies.[89] Our recommendations for improving the design, conduct, analysis, and reporting standards of cancer trials are listed in box 3.[29]

Limitations

Our study had several limitations. We did not include clinical study reports, which, according to previous reviews, might provide the most comprehensive set of information on randomised controlled trials.[50 90] Because there is no established guidance on how to feasibly collect information from such reports, their use in systematic reviews remains limited.[91 92] We focused on cancer drug trials, therefore the generalisability of our findings to trials in other therapeutic areas is unclear. Nevertheless, cancer drugs comprise the single largest category of recent drug approvals.[93 94] Additionally our study included cancer drug approvals between 2014 and 2016; characteristics of randomised controlled trials that supported EMA approvals during our study period might not reflect the design, conduct, analysis, and reporting of cancer drug trials outside of this period. Furthermore, our risk of bias assessments were not blinded to study results because risk of bias assessments require examination of results. However, a systematic review of randomised trials did not identify evidence overall of a difference in risk of bias judgments between blinded and unblinded assessments.[95]

We examined the risk of bias, rather than bias itself. Therefore, it remains a possibility that trial results are unbiased despite the methodological flaws identified in our assessments. According to previous studies, risk of bias judgments based on publications alone might rely on incomplete information,[50] and might not reflect the true methodological rigour of underlying studies.[69 70] To address this issue, we relied on a combination of regulatory documents, trial protocols, publications, supplementary materials, and clinical trials registries. In some cases, our judgments were substantiated with potential evidence of bias; for instance, when outcome measurements available from unblinded local investigators produced exaggerated findings compared with those obtained from an independent panel of masked assessors. For example, the magnitude of progression free survival benefit reported in the BRF113220 trial was less pronounced when assessed by an independent committee than by investigators (hazard ratio 0.39 v 0.55).[96] Notably, our findings do not imply that EMA decisions are biased, and they do not suggest that pharmaceutical manufacturers deliberately introduce bias into their trials. Instead, our findings identify methodological shortcomings in pivotal trials of new cancer drugs.

Finally, our assessments focused on the primary endpoints of randomised controlled trials; it remains possible that results for other outcomes could be at lower risk of bias. However, this is unlikely because pervasive limitations are well documented for secondary endpoints of cancer drug trials, including harms[97-99] and quality of life outcomes.[100 101] Therefore, we might not have fully captured other important shortcomings of randomised controlled trials that support cancer drug approvals.

Conclusions

Most pivotal studies that formed the basis of EMA approval of new cancer drugs between 2014 and 2016 used randomised designs. Despite the widely accepted strengths of such designs, we concluded that almost half of randomised controlled trials were at high risk of bias because of deficits in their design, analysis, and conduct. During our study period, European regulators identified other concerns for 7 out of 16 drugs that had at least one randomised controlled trial at low risk of bias. Journal publications did not acknowledge the key trial limitations identified in regulatory documents. Policymakers, investigators, and clinicians should carefully consider risks of bias in pivotal trials that support regulatory decisions, and the extent to which new cancer therapies offer meaningful benefit to patients.

AUTHOR AFFILIATIONS
[1]Department of Health Policy, London School of Economics and Political Science, London WC2A 2AE, UK

[2]Program on Regulation, Therapeutics, and Law (PORTAL), Division of Pharmacoepidemiology and Pharmacoeconomics, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

[3]Department of Global Health and Social Medicine, King's College London, London, UK

[4]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

[5]The National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care West (NIHR CLAHRC West) at University Hospitals Bristol NHS Foundation Trust, Bristol, UK

[6]National Institute for Health Research Bristol Biomedical Research Centre, Bristol, UK

[7]Cancer Research Institute, Queen's University at Kingston, Kingston, Ontario, Canada

1 The European Parliament and the Council of the European Union. REGULATION (EC) No 726/2004. March 2004. https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2004:136:0001:0033:en:PDF. Accessed on March 4, 2019.

2 Davis C, Abraham J. *Unhealthy Pharmaceutical Regulation: Innovation, Politics and Promissory Science.* Springer, 2013. doi:10.1057/9781137349477

3 Morant AV, Vestergaard HT. European marketing authorizations granted based on a single pivotal clinical trial: the rule or the exception? *Clin Pharmacol Ther* 2018;104:169-77. doi:10.1002/cpt.900

4 European Medicines Agency. Human Medicines Highlight 2017. https://www.ema.europa.eu/en/documents/report/human-medicines-highlights-2017_en.pdf. Accessed January 24, 2019.

5 Tibau A, Molto C, Borrell M, et al. Magnitude of clinical benefit of cancer drugs approved by the US Food and Drug Administration based on single-arm trials. *JAMA Oncol* 2018;4:1610-1. doi:10.1001/jamaoncol.2018.4300

6 Light DW, Lexchin J. Why do cancer drugs get such an easy ride? *BMJ* 2015;350:h2068. doi:10.1136/bmj.h2068

7 Aggarwal A, Fojo T, Chamberlain C, Davis C, Sullivan R. Do patient access schemes for high-cost cancer drugs deliver value to society? Lessons from the NHS Cancer Drugs Fund. *Ann Oncol* 2017;28:1738-50. doi:10.1093/annonc/mdx110

8 Fojo T, Mailankody S, Lo A. Unintended consequences of expensive cancer therapeutics—the pursuit of marginal indications and a me-too mentality that stifles innovation and creativity: the John Conley Lecture. *JAMA Otolaryngol Head Neck Surg* 2014;140:1225-36. doi:10.1001/jamaoto.2014.1570

9 Kumar H, Fojo T, Mailankody S. An appraisal of clinically meaningful outcomes guidelines for oncology clinical trials. *JAMA Oncol* 2016;2:1238-40. doi:10.1001/jamaoncol.2016.0931

10 Gyawali B, Sullivan R. Economics of cancer medicines: for whose benefit? *New Bioeth* 2017;23:95-104. doi:10.1080/20502877.2017.1314885

11 Davis C, Naci H, Gurpinar E, Poplavska E, Pinto A, Aggarwal A. Availability of evidence of benefits on overall survival and quality of life of cancer drugs approved by European Medicines Agency: retrospective cohort study of drug approvals 2009-13. *BMJ* 2017;359:j4530. doi:10.1136/bmj.j4530

12 Booth CM, Eisenhauer EA. Progression-free survival: meaningful or simply measurable? *J Clin Oncol* 2012;30:1030-3. doi:10.1200/JCO.2011.38.7571

13 Prasad V, Kim C, Burotto M, Vandross A. The strength of association between surrogate end points and survival in oncology: a systematic review of trial-level meta-analyses. *JAMA Intern Med* 2015;175:1389-98. doi:10.1001/jamainternmed.2015.2829

14 Del Paggio JC, Azariah B, Sullivan R, et al. Do contemporary randomized controlled trials meet ESMO thresholds for meaningful clinical benefit? *Ann Oncol* 2017;28:157-62.

15 Del Paggio JC, Sullivan R, Schrag D, et al. Delivery of meaningful cancer care: a retrospective cohort study assessing cost and benefit with the ASCO and ESMO frameworks. *Lancet Oncol* 2017;18:887-94. doi:10.1016/S1470-2045(17)30415-1

16 Schumock GT, Li EC, Wiest MD, et al. National trends in prescription drug expenditures and projections for 2017. *Am J Health Syst Pharm* 2017;74:1158-73. doi:10.2146/ajhp170164

17 Tafuri G, Stolk P, Trotta F, et al. How do the EMA and FDA decide which anticancer drugs make it to the market? A comparative qualitative study on decision makers' views. *Ann Oncol* 2014;25:265-9. doi:10.1093/annonc/mdt512

18 Downing NS, Aminawung JA, Shah ND, Krumholz HM, Ross JS. Clinical trial evidence supporting FDA approval of novel therapeutic agents, 2005-2012. *JAMA* 2014;311:368-77. doi:10.1001/jama.2013.282034

19 Kesselheim AS, Myers JA, Avorn J. Characteristics of clinical trials to support approval of orphan vs nonorphan drugs for cancer. *JAMA* 2011;305:2320-6. doi:10.1001/jama.2011.769

20 Jones DS, Podolsky SH. The history and fate of the gold standard. *Lancet* 2015;385:1502-3. doi:10.1016/S0140-6736(15)60742-5

21 Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408-12. doi:10.1001/jama.1995.03520290060030

22 Wood L, Egger M, Gluud LL, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ* 2008;336:601-5. doi:10.1136/bmj.39465.451744.AD

23 Hróbjartsson A, Forfang E, Haahr MT, Als-Nielsen B, Brorson S. Blinded trials taken to the test: an analysis of randomized clinical trials that report tests for the success of blinding. *Int J Epidemiol* 2007;36:654-63. doi:10.1093/ije/dym020

24 Nüesch E, Trelle S, Reichenbach S, et al. The effects of excluding patients from the analysis in randomised controlled trials: meta-epidemiological study. *BMJ* 2009;339:b3244. doi:10.1136/bmj.b3244

25 Abraha I, Cherubini A, Cozzolino F, et al. Deviation from intention to treat analysis in randomised trials and treatment effect estimates: meta-epidemiological study. *BMJ* 2015;350:h2445. doi:10.1136/bmj.h2445

26 Page MJ, Higgins JPT, Clayton G, Sterne JAC, Hróbjartsson A, Savović J. Empirical evidence of study design biases in randomized trials: systematic review of meta-epidemiological studies. *PLoS One* 2016;11:e0159267. doi:10.1371/journal.pone.0159267

27 Savović J, Jones HE, Altman DG, et al. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. *Ann Intern Med* 2012;157:429-38. doi:10.7326/0003-4819-157-6-201209180-00537

28 Higgins JPT, Altman DG, Gøtzsche PC, et al, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928. doi:10.1136/bmj.d5928

29 Booth CM, Cescon DW, Wang L, Tannock IF, Krzyzanowska MK. Evolution of the randomized controlled trial in oncology over three decades. *J Clin Oncol* 2008;26:5458-64. doi:10.1200/JCO.2008.16.5456

30 Fleming TR, Rothmann MD, Lu HL. Issues in using progression-free survival when evaluating oncology products. *J Clin Oncol* 2009;27:2874-80. doi:10.1200/JCO.2008.20.4107

31 Naci H, Smalley KR, Kesselheim AS. Characteristics of preapproval and postapproval studies for drugs granted accelerated approval by the US Food and Drug Administration. *JAMA* 2017;318:626-36. doi:10.1001/jama.2017.9415

32 Mansournia MA, Higgins JPT, Sterne JAC, Hernán MA. Biases in randomized trials: a conversation between trialists and epidemiologists. *Epidemiology* 2017;28:54-9. doi:10.1097/EDE.0000000000000564

33 Dechartres A, Trinquart L, Atal I, et al. Evolution of poor reporting and inadequate methods over time in 20 920 randomised controlled trials included in Cochrane reviews: research on research study [correction in: *BMJ* 2017;358:j3806]. *BMJ* 2017;357:j2490. doi:10.1136/bmj.j2490

34 Köhler M, Haag S, Biester K, et al. Information on new drugs at market entry: retrospective analysis of health technology assessment reports versus regulatory reports, journal publications, and registry reports. *BMJ* 2015;350:h796. doi:10.1136/bmj.h796

35 Hoekman J, Boon W. Changing standards for drug approval: a longitudinal analysis of conditional marketing authorisation in the European Union. *Soc Sci Med* 2019;222:76-83. doi:10.1016/j.socscimed.2018.12.025

36 Banzi R, Gerardi C, Bertele' V, Garattini S. Approvals of drugs with uncertain benefit-risk profiles in Europe. *Eur J Intern Med* 2015;26:572-84. doi:10.1016/j.ejim.2015.08.008

37 Banzi R, Gerardi C, Bertele' V, Garattini S. Conditional approval of medicines by the EMA. *BMJ* 2017;357:j2062. doi:10.1136/bmj.j2062

38 Tse T, Fain KM, Zarin DA. How to avoid common problems when using ClinicalTrials.gov in research: 10 issues to consider. *BMJ* 2018;361:k1452. doi:10.1136/bmj.k1452

39 Higgins JPT, Sterne JAC, Savović J, et al. A revised tool for assessing risk of bias in randomized trials. *Cochrane Database Syst Rev* 2016;10(Suppl 1):29-31.

40 Page MJ, Shamseer L, Altman DG, et al. Epidemiology and reporting characteristics of systematic reviews of biomedical research: a cross-sectional study. *PLoS Med* 2016;13:e1002028. doi:10.1371/journal.pmed.1002028

41 Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:l4898. doi:10.1136/bmj.l4898

42 Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. *Lancet* 2002;359:614-8. doi:10.1016/S0140-6736(02)07750-4

43 Clark L, Fairhurst C, Torgerson DJ. Allocation concealment in randomised controlled trials: are we getting better? *BMJ* 2016;355:i5663. doi:10.1136/bmj.i5663

44 Bello S, Moustgaard H, Hróbjartsson A. The risk of unblinding was infrequently and incompletely reported in 300 randomized clinical trial publications. *J Clin Epidemiol* 2014;67:1059-69. doi:10.1016/j.jclinepi.2014.05.007

45 Bello S, Moustgaard H, Hróbjartsson A. Unreported formal assessment of unblinding occurred in 4 of 10 randomized clinical trials, unreported loss of blinding in 1 of 10 trials. *J Clin Epidemiol* 2017;81:42-50. doi:10.1016/j.jclinepi.2016.08.002

46 Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;340:c869. doi:10.1136/bmj.c869

47 Sridhara R, Mandrekar SJ, Dodd LE. Missing data and measurement variability in assessing progression-free survival endpoint in

randomized clinical trials. *Clin Cancer Res* 2013;19:2613-20. doi:10.1158/1078-0432.CCR-12-2938

48 Bhattacharya S, Fyfe G, Gray RJ, Sargent DJ. Role of sensitivity analyses in assessing progression-free survival in late-stage oncology trials. *J Clin Oncol* 2009;27:5958-64. doi:10.1200/JCO.2009.22.4329

49 Hróbjartsson A, Thomsen ASS, Emanuelsson F, et al. Observer bias in randomised clinical trials with binary outcomes: systematic review of trials with both blinded and non-blinded outcome assessors. *BMJ* 2012;344:e1119. doi:10.1136/bmj.e1119

50 Wieseler B, Kerekes MF, Vervoelgyi V, McGauran N, Kaiser T. Impact of document type on reporting quality of clinical drug trials: a comparison of registry reports, clinical study reports, and journal publications. *BMJ* 2012;344:d8141. doi:10.1136/bmj.d8141

51 Woloshin S, Schwartz LM, Bagley PJ, Blunt HB, White B. Characteristics of interim publications of randomized clinical trials and comparison with final publications. *JAMA* 2018;319:404-6. doi:10.1001/jama.2017.20653

52 Goodman SN. Stopping at nothing? Some dilemmas of data monitoring in clinical trials. *Ann Intern Med* 2007;146:882-7. doi:10.7326/0003-4819-146-12-200706190-00010

53 Hatswell AJ, Baio G, Berlin JA, Irs A, Freemantle N. Regulatory approval of pharmaceuticals without a randomised controlled study: analysis of EMA and FDA approvals 1999-2014. *BMJ Open* 2016;6:e011666. doi:10.1136/bmjopen-2016-011666

54 European Medicines Agency. EMA/CHMP/205/95 Rev.5: Guideline on the evaluation of anticancer medicinal products. September 2017. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-evaluation-anticancer-medicinal-products-man-revision-5_en.pdf. Accessed January 29, 2019.

55 Haslam A, Hey SP, Gill J, Prasad V. 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* 2019;106:196-211. doi:10.1016/j.ejca.2018.11.012

56 Kovic B, Jin X, Kennedy SA, et al. Evaluating progression-free survival as a surrogate outcome for health-related quality of life in oncology: a systematic review and quantitative analysis. *JAMA Intern Med* 2018;178:1586-96. doi:10.1001/jamainternmed.2018.4710

57 Amir E, Seruga B, Kwong R, Tannock IF, Ocaña A. Poor correlation between progression-free and overall survival in modern clinical trials: are composite endpoints the answer? *Eur J Cancer* 2012;48:385-8. doi:10.1016/j.ejca.2011.10.028

58 Hwang TJ, Gyawali B. Association between progression-free survival and patients' quality of life in cancer clinical trials. *Int J Cancer* 2019;144:1746-51. doi:10.1002/ijc.31957

59 Kumar S, Rajkumar SV. Surrogate endpoints in randomised controlled trials: a reality check. *Lancet* 2019;394:281-3. doi:10.1016/S0140-6736(19)31711-8.

60 Prasad V, Bilal U. The role of censoring on progression free survival: oncologist discretion advised. *Eur J Cancer* 2015;51:2269-71. doi:10.1016/j.ejca.2015.07.005

61 Templeton AJ, Ace O, Amir E, et al. Influence of censoring on conclusions of trials for women with metastatic breast cancer. *Eur J Cancer* 2015;51:721-4. doi:10.1016/j.ejca.2014.12.016

62 Pignatti F, Hemmings R, Jonsson B. Is it time to abandon complete blinded independent central radiological evaluation of progression in registration trials? *Eur J Cancer* 2011;47:1759-62. doi:10.1016/j.ejca.2011.05.009

63 Food and Drug Administration. *Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics Guidance for Industry.* https://www.fda.gov/media/71195/download. Accessed May 20, 2019.

64 Corbett MS, Higgins JPT, Woolacott NF. Assessing baseline imbalance in randomised trials: implications for the Cochrane risk of bias tool. *Res Synth Methods* 2014;5:79-85. doi:10.1002/jrsm.1090

65 Booth CM, Tannock I. Reflections on medical oncology: 25 years of clinical trials—where have we come and where are we going? *J Clin Oncol* 2008;26:6-8. doi:10.1200/JCO.2007.13.8156

66 Kordecka A, Walkiewicz-Żarek E, Łapa J, Sadowska E, Kordecki M. Selection of endpoints in clinical trials: trends in European marketing authorization practice in oncological indications. *Value Health* 2019;22:884-90. doi:10.1016/j.jval.2019.03.007

67 Reveiz L, Chapman E, Asial S, Munoz S, Bonfill X, Alonso-Coello P. Risk of bias of randomized trials over time. *J Clin Epidemiol* 2015;68:1036-45. doi:10.1016/j.jclinepi.2014.06.001

68 Vale CL, Tierney JF, Burdett S. Can trial quality be reliably assessed from published reports of cancer trials: evaluation of risk of bias assessments in systematic reviews. *BMJ* 2013;346:f1798. doi:10.1136/bmj.f1798

69 Mhaskar R, Djulbegovic B, Magazin A, Soares HP, Kumar A. Published methodological quality of randomized controlled trials does not reflect the actual quality assessed in protocols. *J Clin Epidemiol* 2012;65:602-9. doi:10.1016/j.jclinepi.2011.10.016

70 Soares HP, Daniels S, Kumar A, et al, Radiation Therapy Oncology Group. Bad reporting does not mean bad methods for randomised trials: observational study of randomised controlled trials performed by the Radiation Therapy Oncology Group. *BMJ* 2004;328:22-4. doi:10.1136/bmj.328.7430.22

71 Jefferson T, Jones MA, Doshi P, et al. Risk of bias in industry-funded oseltamivir trials: comparison of core reports versus full clinical study reports. *BMJ Open* 2014;4:e005253. doi:10.1136/bmjopen-2014-005253

72 Tannock IF, Amir E, Booth CM, et al. Relevance of randomised controlled trials in oncology. *Lancet Oncol* 2016;17:e560-7. doi:10.1016/S1470-2045(16)30572-1

73 Naci H, Wouters OJ, Gupta R, Ioannidis JPA. Timing and characteristics of cumulative evidence available on novel therapeutic agents receiving Food and Drug Administration accelerated approval. *Milbank Q* 2017;95:261-90. doi:10.1111/1468-0009.12261

74 Naci H, Carter AW, Mossialos E. Why the drug development pipeline is not delivering better medicines. *BMJ* 2015;351:h5542. doi:10.1136/bmj.h5542

75 Sorenson C, Naci H, Cylus J, Mossialos E. Evidence of comparative efficacy should have a formal role in European drug approvals. *BMJ* 2011;343:d4849. doi:10.1136/bmj.d4849

76 Bell ML, Kenward MG, Fairclough DL, Horton NJ. Differential dropout and bias in randomised controlled trials: when it matters and when it may not. *BMJ* 2013;346:e8668. doi:10.1136/bmj.e8668

77 Boutron I, Estellat C, Guittet L, et al. Methods of blinding in reports of randomized controlled trials assessing pharmacologic treatments: a systematic review. *PLoS Med* 2006;3:e425. doi:10.1371/journal.pmed.0030425

78 Yordanov Y, Dechartres A, Porcher R, Boutron I, Altman DG, Ravaud P. Avoidable waste of research related to inadequate methods in clinical trials. *BMJ* 2015;350:h809. doi:10.1136/bmj.h809

79 Zhang J, Zhang Y, Tang S, et al. 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* 2018;8:e017240. doi:10.1136/bmjopen-2017-017240

80 Hróbjartsson A, Thomsen ASS, Emanuelsson F, et al. Observer bias in randomized clinical trials with time-to-event outcomes: systematic review of trials with both blinded and non-blinded outcome assessors. *Int J Epidemiol* 2014;43:937-48. doi:10.1093/ije/dyt270

81 Chen EY, Joshi SK, Tran A, Prasad V. Estimation of study time reduction using surrogate end points rather than overall survival in oncology clinical trials. *JAMA Intern Med* 2019;179:642-7. doi:10.1001/jamainternmed.2018.8351

82 Jenkins V, Farewell V, May S, et al. Do drugs offering only PFS maintain quality of life sufficiently from a patient's perspective? Results from AVALPROFS (Assessing the 'VALue' to patients of PROgression Free Survival) study. *Support Care Cancer* 2018;26:3941-9. doi:10.1007/s00520-018-4273-3

83 Fallowfield LJ, Catt SL, May SF, et al. Therapeutic aims of drugs offering only progression-free survival are misunderstood by patients, and oncologists may be overly optimistic about likely benefits. *Support Care Cancer* 2017;25:237-44. doi:10.1007/s00520-016-3408-7

84 Lenzer J. FDA committee votes to withdraw bevacizumab for breast cancer. *BMJ* 2011;343:d4244. doi:10.1136/bmj.d4244

85 Kim C, Prasad V. Cancer drugs approved on the basis of a surrogate end point and subsequent overall survival: an analysis of 5 years of US Food and Drug Administration Approvals. *JAMA Intern Med* 2015;175:1992-4. doi:10.1001/jamainternmed.2015.5868

86 Turner L, Shamseer L, Altman DG, Schulz KF, Moher D. Does use of the CONSORT Statement impact the completeness of reporting of randomised controlled trials published in medical journals? A Cochrane review. *Syst Rev* 2012;1:60. doi:10.1186/2046-4053-1-60

87 Berntgen M, Gourvil A, Pavlovic M, Goettsch W, Eichler H-G, Kristensen FB. Improving the contribution of regulatory assessment reports to health technology assessments--a collaboration between the European Medicines Agency and the European network for Health Technology Assessment. *Value Health* 2014;17:634-41. doi:10.1016/j.jval.2014.04.006

88 Barbui C, Baschirotto C, Cipriani A. EMA must improve the quality of its clinical trial reports. *BMJ* 2011;342:d2291. doi:10.1136/bmj.d2291

89 Doshi P, Godlee F. The wider role of regulatory scientists. *BMJ* 2017;357:j1991. doi:10.1136/bmj.j1991

90 Wieseler B, Wolfram N, McGauran N, et al. Completeness of reporting of patient-relevant clinical trial outcomes: comparison of unpublished clinical study reports with publicly available data. *PLoS Med* 2013;10:e1001526. doi:10.1371/journal.pmed.1001526

91 Schroll JB, Bero L, Gøtzsche PC. Searching for unpublished data for Cochrane reviews: cross sectional study. *BMJ* 2013;346:f2231. doi:10.1136/bmj.f2231

92 Hodkinson A, Dietz KC, Lefebvre C, et al. The use of clinical study reports to enhance the quality of systematic reviews: a survey of systematic review authors. *Syst Rev* 2018;7:117. doi:10.1186/s13643-018-0766-x

93 Mullard A. 2017 FDA drug approvals [correction in: *Nat Rev Drug Discov* 2018;17:150]. *Nat Rev Drug Discov* 2018;17:81-5. doi:10.1038/nrd.2018.4

94 Mullard A. EMA recommended 35 new drugs in 2017. *Nat Rev Drug Discov* 2018;17:86. doi:10.1038/nrd.2018.17

95 Morissette K, Tricco AC, Horsley T, Chen MH, Moher D. Blinded versus unblinded assessments of risk of bias in studies included in a systematic review. *Cochrane Database Syst Rev* 2011;(9):MR000025. doi:10.1002/14651858.MR000025.pub2

96 Flaherty KT, Infante JR, Daud A, et al. Combined BRAF and MEK inhibition in melanoma with BRAF V600 mutations. *N Engl J Med* 2012;367:1694-703. doi:10.1056/NEJMoa1210093

97 Vera-Badillo FE, Shapiro R, Ocana A, Amir E, Tannock IF. Bias in reporting of end points of efficacy and toxicity in randomized, clinical trials for women with breast cancer. *Ann Oncol* 2013;24:1238-44. doi:10.1093/annonc/mds636

98 Vera-Badillo FE, Napoleone M, Krzyzanowska MK, et al. Bias in reporting of randomised clinical trials in oncology. *Eur J Cancer* 2016;61:29-35. doi:10.1016/j.ejca.2016.03.066

99 Sivendran S, Latif A, McBride RB, et al. Adverse event reporting in cancer clinical trial publications. *J Clin Oncol* 2014;32:83-9. doi:10.1200/JCO.2013.52.2219

100 Bylicki O, Gan HK, Joly F, Maillet D, You B, Péron J. Poor patient-reported outcomes reporting according to CONSORT guidelines in randomized clinical trials evaluating systemic cancer therapy. *Ann Oncol* 2015;26:231-7. doi:10.1093/annonc/mdu489

101 Gyawali B, Shimokata T, Honda K, Ando Y. Reporting harms more transparently in trials of cancer drugs [correction in: *BMJ* 2018;363:k4930]. *BMJ* 2018;363:k4383. doi:10.1136/bmj.k4383

Web appendix: Supplementary material