Strength of Evidence Supporting Cancer Drug Approvals in China, 2017-2021: A Cross-Sectional Analysis

Yichen Zhang, Dingyi Chen, Mengyuan Fu, Luwen Shi, Huseyin Naci, Anita K. Wagner, Joseph S. Ross, Xiaodong Guan

Affiliations:

Department of Pharmacy Administration and Clinical Pharmacy, School of Pharmaceutical Sciences, Peking University, Beijing, China (Y Zhang PhD, D Chen MSc, Prof L Shi MSc, X Guan PhD)

International Research Centre for Medicinal Administration, Peking University, Beijing, China (M Fu PhD, Prof L Shi MSc, X Guan PhD)

Section of General Medicine, Department of Medicine, Yale School of Medicine, New Haven, Connecticut, USA (M Fu PhD, Prof J S Ross MD MHS)

LSE Health, Department of Health Policy, London School of Economics and Political Science, London, UK (H Naci PhD)

Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, Massachusetts, USA (A K Wagner DrPH)

Correspondence to: Xiaodong Guan, Department of Pharmacy Administration and Clinical Pharmacy, School of Pharmaceutical Sciences, Peking University, Beijing 100191, China, **guanxiaodong@pku.edu.cn**

Original Research

Manuscript words: 3,504 (excluding title page, abstract, references, figures and tables)

References: 31

Tables: 2

Figures: 4

Appendix: 1

Summary

Background

Well-designed pivotal clinical trials can provide robust evidence for the market authorization of new cancer drugs, whereas lower-quality clinical evidence leads to uncertainty about drug benefits and harms. We aim to investigate the strength of evidence supporting new cancer drug indications approved in China from 2017 to 2021.

Methods

In this cross-sectional analysis, we searched publicly available data from the National Medical Products Administration website to identify pivotal pre-approval efficacy trials supporting all original and supplemental cancer drug indications approved in China from 2017 to 2021. We collected trial protocols and publications from ClinicalTrials.gov, PubMed, and the China National Knowledge Infrastructure database. The primary outcome was the strength of the supporting pivotal studies, as measured by study design (randomized or single-arm) and quality. For study quality, we evaluated the ability to minimize bias of single-arm trials, measured as adopted external control arm and adjusted confounders; risk of bias of randomized controlled trials (RCTs), as evaluated using the revised Cochrane tool for risk of bias assessment. Additionally, we used ratio of hazard ratios (RHR) to quantify differences in effect size in RCTs with different risks of bias.

Findings

Between 1 January 2017 and 31 December 2021, 77 novel cancer drugs for 86 original and 62 supplemental indications were approved in China, based on data from 205 pivotal studies. Forty-four (29·7%) indications were supported by single-arm trials only, and 104 (70·3%) indications were supported by at least one RCT. Of the 56 pivotal single-arm trials with regulatory review documents, 6 (10·7%) used aggregated data from earlier trials as external controls, without adjustment for confounders. Of the 128 pivotal RCTs with published results, 47 (36·7%) and 48 (37·5%) were assessed as having some concern or a high risk of bias, respectively. RCTs judged to be at some concern or high risk of bias in the randomization process had smaller effect sizes (RHR=0·678, 95% confidence interval: 0·532 to 0·864), and those judged to be at some concern or high risk of bias in

missing outcome data had larger effect sizes (RHR=1·114, 95% confidence interval: 1·004 to 1·237), compared to RCTs at low risk of bias in these domains.

Interpretation

Four-fifths of assessable pivotal studies supporting new cancer indication approvals in China from 2017-2021 had weaknesses in design, conduct, or reporting that introduce uncertainty to the estimation of treatment effects. To ensure the validity of drug efficacy data and reduce uncertainty, stakeholders should strengthen and implement a high-quality standard on the design, conduct, analysis, and reporting of studies supporting regulatory approval of new therapies.

Funding

National Natural Science Foundation of China (Grant No. 72274004).

Research in context

Evidence before this study

Over the past decade, a series of regulatory guidance concerning drug clinical trial design, analysis, and implementation has been issued in China. Randomized controlled trials (RCTs) are still considered the gold standard for evaluating the efficacy and safety of a drug. For cancer drug trials pursuing regulatory approval, an evaluation is suggested to address the uncertainty and bias related to the assessment of clinical benefit. We searched PubMed for peer-reviewed, original studies (from database inception to Jan 31, 2025), using the search terms “strength of evidence”, “evidence quality”, “trial assessment”, and “cancer drug”. Several observational studies conducted in the US characterized the clinical trials of drugs granted regular and accelerated approval. Some research examined the design characteristics and risk of bias of pivotal RCTs of cancer drugs approved by the European Medicines Agency. Our previous study evaluated the quality of control arms of investigational new drugs in China. No literature assessed the biases in cancer drug pivotal trials in China.

Added value of this study?

We report the strength of evidence supporting new cancer drug indications approved in China from 2017 to 2021. Approximately one-third of cancer drug indication approvals were supported by single-arm trials. While two-thirds were supported by RCTs, over one-third of these RCTs were judged to have a high risk of bias. Overall, fewer than one-fifth of the approved indications were based on RCTs with a low risk of bias. The exploratory meta-epidemiology analysis showed that, compared to the RCTs at low risk of bias, those with some concern or a high risk of bias in missing outcome data showed larger effect sizes, principally in those with surrogate primary endpoints.

Implications of all the available evidence

A more robust premarketing evaluation of investigational new cancer drugs is needed to improve certainty regarding drug benefits and harms in China, as a lack of strong efficacy evidence may distort the appraisal of drug clinical benefits, risk-benefit evaluations, and further challenge the pricing and reimbursement of expensive cancer therapies.

Introduction

Well-designed and adequately conducted clinical trials are the cornerstone to demonstrating drug safety and efficacy and support regulatory approval of drugs.(1) In general, applications for market authorization of a drug for a new indication should be supported by two rigorously-designed randomized controlled trials (RCTs) that demonstrate the drug's clinical benefits, or one adequate and well-controlled large multicenter trial that can provide substantial evidence of effectiveness.(2, 3) Yet, regulators increasingly make approval decisions based on uncertain or insufficient evidence of benefit, especially for drugs indicated for life-threatening diseases like cancers.(4) In 2000-2020, fifty percent of U.S. Food and Drug Administration (FDA)-approved cancer indications were authorized in the absence of RCT evidence.(5) In Europe, around half of RCTs supporting approvals of cancer drugs by the European Medicines Agency (EMA) between 2014 and 2016 were judged to be at high risk of bias,(6) defined as the likelihood that features of the study design, conduct, analysis or reporting of the trial will lead to systematic error or deviation from the truth in results or inferences.(7)

The strength of evidence underpinning regulatory approval of new drugs has changed over the past decades, due to more flexibility in regulatory standards.(4) In China, which represents one of the largest global pharmaceutical markets, regulatory reforms that started in 2015 incentivized the growth of new drug research and development.(8) A series of subsequent technical guidelines concerning the design of pivotal investigational new drug trials recommended RCTs, as random allocation can reduce selection bias and thereby enhance the internal validity.(9) For indications for which implementing RCTs is not feasible, Chinese regulatory authorities also emphasized that the adoption of single-arm trials as a pivotal study was acceptable but should be implemented with caution, as such designs would introduce substantial uncertainty in the drug risk-benefit assessment (**appendix p 1**).

Cancer is a growing public health problem in China. Over the past two decades, China has also issued strategies to incentivize cancer drug research and development, in conjunction with regulatory

reforms to improve the availability of new cancer drugs.(10) Since 2017, the number of cancer drug approvals, especially those developed by domestic pharmaceutical companies, has increased in China.(10) Earlier research provided a preliminary characterization of clinical trials supporting their original market authorization.(11) To date, however, no study has investigated the strength and quality of evidence supporting cancer drug indication approvals in China, including sources of bias. Bias in clinical trials, defined as the systematic error, or deviation from the truth, in results, will lead to under-estimation or over-estimation of the true intervention effect and can vary in magnitude.(7) Accordingly, we aimed to examine the strength of evidence of the pivotal efficacy studies supporting cancer drug approvals in China and to assess their design and quality, measured by the risk of bias in RCTs and the ability to adjust for confounders in single-arm trials.

Methods

Study design and data sources

In this cross-sectional study, we identified all cancer drugs and corresponding indications approved in China between 1 January 2017 and 31 December 2021, including both original and supplemental indication approvals, as described previously.(10) Cancer drug authorizations were identified using the quarterly *National Drug Code Data File* issued by the National Medical Products Administration (NMPA) which contained all medical products available on the Chinese market.(12) We included small molecules as well as biologics, and excluded traditional Chinese medicines, prophylactic vaccines, and generic or biosimilar versions of previously approved drugs. Cancer drugs were categorized as those authorized in China only and those also authorized by the FDA or the EMA (**appendix p 3**). We also recorded whether each indication received regular or conditional marketing authorization. For each drug, we assessed the review documents and product labels to identify approved indications for adult malignancies. We included both original and supplemental indications, as reported in the regulatory review documents. Indications for pediatric use only, benign tumors, and supportive care were excluded (**Figure 1**). All cancer drug indications were categorized into first-line, later-line, adjuvant or neoadjuvant, and maintenance treatments.

For each indication, we identified the pivotal studies (i.e., pivotal efficacy trials, including RCTs, single-arm trials, and dose-optimization trials) described in the "Effectiveness Evaluation" section of the review documents issued by the Center for Drug Evaluation, NMPA.(13) For each pivotal study, we collected the study name and identifiers then crosschecked the corresponding clinical trial identifiers in ClinicalTrials.gov and chinadrugstrials.org.cn (official trial registry of the NMPA)

Using clinical trial identifiers, we searched PubMed and ClinicalTrials.gov to retrieve any associated peer-reviewed publications and protocols of pivotal studies published by 31 December 2023, to allow for a minimum follow-up duration of two years since approval. If no record was available, we further searched trial names in combination with approved indications in PubMed and the China National Knowledge Infrastructure Database (one of the most commonly used Chinese literature datasets, **appendix p 3**). Only publications reporting the primary efficacy endpoint results were included.

Using trial publications, we extracted information on trial features including design, randomization, phase, region of enrolment, endpoint(s), sample size, and comparator.(14) When trial publications were not publicly available, we relied on information from clinicaltrials.gov and/or chinadrugstrials.org.cn. Trial designs were categorized as randomized clinical trials or single-arm trials, including four dose-response trials. For RCTs with published results, we also collected reported estimates of effect sizes.

Procedures

To assess the ability to minimize bias of single-arm trials, for cancer drug indications supported by single-arm trials only, we reviewed their regulatory review documents to check whether non-concurrent controls (i.e., historical control or data collected outside the study contemporaneously) were adopted.(15) For studies comparing efficacy data with those of prior research, we further examined whether they clarified the specific time of historical data, as disease diagnostic criteria and

efficacy evaluation methods may change with the development of medical practice.(16) We also examined whether they adjusted for any confounders in the comparison.(14-16)

For RCTs, we used the Cochrane revised tool for assessing risk of bias (RoB 2, the 22 August 2019 version) to assess their risk of bias (**appendix p 4**).(17) Two investigators (D.C. and Y.Z.) independently assessed the risk of bias (i.e., low, some concern, or high) in each trial. The interrater reliability for the first round of assessment was moderate (Cohen's kappa=0·509, agreement rate=68·8%). We found disagreement was predominantly in the third domain (missing outcome data) of RoB2 and unified our evaluation criteria (**appendix p 4**). The second round of interrater reliability improved (Cohen's kappa= 0·795, agreement rate=86·6%). Discrepancies were solved by consensus. This study was considered not involving human subjects by the Institutional Review Board of Peking University.

Outcomes

The primary outcome was the strength of the supporting pivotal studies, as measured by study design (randomized or single-arm) and quality (adopting external control arms and adjusted confounders of single-arm trial, and risk of bias of RCTs). The secondary outcome was the differences in effect size in RCTs with different risks of bias, as measured by ratio of hazard ratios (RHR).

Statistical analysis

We descriptively characterized features of cancer drugs, indications, and pivotal trials. We also counted the number of pivotal studies and pivotal RCTs with a low risk of bias per indication over time. We examined the annual trend in the proportion of indications with at least one low-risk RCT using the Cochran-Armitage trend test. Univariable logistic regression was performed to examine the association between risk of bias judgments and RCT features (i.e., the primary endpoint, line-of-therapy, trial location, trial comparator), as multivariable regression does not improve the model fit (**appendix p 6**). Two-sided P values of $< 0{\cdot}05$ were considered statistically significant.

In an exploratory analysis, we used meta-regression to investigate the relationship between different risk of bias domains and effect size for RCTs reporting time-to-event outcomes.(18, 19) We did not explore the association between the selective outcome reporting domain and effect estimates, in line with a previous study.(19) Considering potential correlations among risk of bias domains, we included all other four domains simultaneously in the regression model. To control for potential confounding, we performed mixed-effects meta-regressions (Maximum Likelihood method) adjusting for cancer site, trial endpoint, trial location, and trial comparator. The analysis estimated RHRs comparing effect sizes in trials with high or some concerns of risk of bias versus low risk of bias.(20) As an HR less than 1·0 indicated a beneficial effect of the experimental intervention, RHRs less than 1·0 implied smaller effect size (i.e., HR) and greater treatment effects (i.e., ability to decrease death or disease progress risk) in trials with a high or some concern risk of bias.(21) Subgroup analyses and between-group homogeneity tests were performed by group with ten or more trials.(22) According to the subgroup analysis results, we performed meta-regression for overall survival (OS) and surrogate endpoints, separately. All modeling analyses were performed in R (version 4.3.3).

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

Between 1 January 2017 and 31 December 2021, 77 novel cancer drugs received marketing authorization in China. Of the 148 corresponding cancer drug indication approvals, 143 (96·6%) had publicly available review documents, whereas 5 (3·4%) did not (**Table 1, appendix p 7**). Among the 205 pivotal studies, 135 (65·9%) were RCTs and 70 (34·1%) were single-arm trials (**Table 2, appendix p 8**). Of the 148 cancer drug indications, 104 (70·3%) were supported by at least one RCT, whereas 44 (29·7%) were supported by single-arm trials only (**Table 1)**. From 2017 to 2021, the proportion of indications supported by at least one RCT decreased from 75·0% (12/16) to 40·7%

(11/27, **Fig 2A**). Fifty one (59·3%) of 86 original indications and 53 (85·5%) of 62 supplementary indications were supported by RCTs, respectively **(appendix p 23**).

Of the 205 pivotal studies, 184 (89·8%) had publication reporting pre-planned results, while only 137 (66·8%) had publicly available trial protocols. Among the 135 RCTs, 128 (94·8%, supporting approvals of 102 indications) had corresponding publications that were publicly available for risk of bias evaluation (**appendix p 8**). Overall, of the 128 assessable RCTs, 48 (37·5%) were assessed as having high risk of bias, most often due to missing outcome data (n=46), while 47 (36·7%) were evaluated as having some concerns, most often due to the deviation from intended interventions (n=59), selection of the reported results (n=39), randomization process (n=7), and measurement of the outcome data (n=3); 33 (25·8%) were assessed as having low risk of bias (**Fig 3A**, **appendix p 28**). Among the 148 indication approvals, 29 (19·6%) were supported by at least one RCT that was assessed as having low risk of bias, and the yearly proportion was 28·6% (n=10/42) in 2018 and 14·8% (n=4/27) in 2021 (p=0·88, **Fig 2B**).

The proportion of RCTs with some concern or high risk of bias is higher in trials for international multicenter trial than in regional multicenter trials (79·5% vs. 62·5%, Odds Ratio [OR]=0·384, 95% Confidence Interval [CI]=0·169 to 0·864, **appendix p 24**). There was no difference in the likelihood of the RCTs being assessed as having some concern or high risk of bias between those supporting cancer drugs authorized in China only (17 of 27 [63·0%]) and those also authorized by the FDA or EMA (78 of 101 [77·2%]; OR=0·465, 95%CI=0·190 to 1·165), and those supporting regular approvals (79 of 109 [72·5%]) and conditional approvals (16 of 19 [84·2%; OR=1·809, 95%CI=0.·584 to 7·282]).

For 44 indications supported by single-arm trials only, **Fig 3B** shows the proportion using external non-concurrent controls. In total, there were 54 (96·4%) of 56 single-arm trials with publicly available regulatory review documents, while the other two trials did not. Of these, 22 (39·3%) did not clarify in the regulatory review documents whether non-concurrent controls were used for their

approvals, and 26 (46·4%) reported using a historical control without clarifying the specific time. Only 6 (10·7%) of 56 pivotal single-arm trials with external controls had specified periods, with a median time of 2·2 years (interquartile range: -0·3 to 5·5 years) between the start of external control studies and pivotal studies (**appendix p 25**). Regulatory review documents did not report adjustments for confounders for any single-arm trials.

Among 122 pivotal RCTs that reported hazard ratio for time-to-event endpoints, 45 (36·9%) used OS as the primary endpoint, and 77 (63·1%) used surrogate endpoints. In the multivariable meta-regression models, different domains of risk of bias, cancer site, line of therapy, trial comparator, and trial location were included (**appendix p 24**). The regression model's coefficient of determination (R-Square) was 82·67%.

For all 122 RCTs reporting time-to-event outcomes, those with some concern risk of bias in the randomization process had a larger treatment effect than those with low risk of bias (RHR=0·678, 95% CI: 0·532 to 0·864), and those with some concern risk of bias in missing outcome data had a smaller treatment effect (RHR=1·114, 95% CI: 1·004 to 1·237, **Fig 4A**). In meta-regression models including an interaction term between risk of domain and other subgroups, we observed a statistically significant interaction between the randomization process domain and endpoint type (p = 0·049), and the missing outcome data domain and endpoint type (p = 0·040), but statistically non-significant differences for trial location and cancer site subgroup (all p>0·05). Studies that reported surrogate outcomes rather than overall survival typically had larger treatment effects (e.g., RHR for progression-free survival vs OS=0·785, 95% CI 0·688 to 0·894, **appendix p 26**). In RCTs that used OS as the primary endpoint, no statistically significant association was observed between risk of bias and over- or underestimation of effect size (**Fig 4B**). In RCTs that used surrogate endpoints as the primary endpoint, compared with trials assessed at low risk of bias, those with some concern or high risk of bias in the randomization process had a larger treatment effect (RHR=0·713, 95% CI: 0·523 to 0·974); those with some concern or high risk of bias in missing outcome data had a smaller treatment effect (RHR=1·173, 95% CI: 1·003 to 1·371, **Fig 4C**).

Discussion

We found that approximately one-third of China's cancer drug indication approvals between 2017 and 2021 were supported by single-arm trials. While two-thirds were supported by RCTs, over one-third of these RCTs were judged to have a high risk of bias. Overall, fewer than 20% of the approved indications were based on RCTs with a low risk of bias. Compared to those evaluated as low risk of bias, RCTs judged to be at some concern or high risk of bias in missing outcome data showed smaller treatment effects, principally in those with surrogate primary endpoints.

This study provides systematic evidence on the strength and quality of pivotal studies underpinning cancer drug approvals following regulatory reforms aimed at accelerating patient access to new drugs in China. Previous research during a similar time horizon showed that half of the trials supporting initial cancer drug approvals in China were non-randomized.(11) Using the Cochrane risk of bias tool and evaluation of the ability to minimize bias, our study further expanded current knowledge on the pivotal study quality and potential biases. The meta-epidemiological exploration of the association between trial risk of bias and effect estimation complements recent research about the trial design and outcomes among FDA-approved cancer drug therapies.(23) Although direct comparisons with the assessment of cancer drug pivotal studies in the US and EU are limited by differing timeframes, our findings are relevant in a broader global regulatory context, where the FDA or EMA approved more new cancer drugs faster based on fewer pivotal trials or less rigorous designs.(4-6, 24)

Although China's 2012 regulatory guidance on cancer drug clinical trials recommended the inclusion of two adequate and well-controlled studies for new drug approval,(1) our results show that this recommendation was not implemented consistently, as only 28·4% of approvals were supported by two or more pivotal studies and just 19·6% included at least one RCT with low risk of bias. This could, in some way, be attributed to flexible regulatory standards to expedite drug access for patients with unmet needs,(4, 25) and the practical challenge of conducting large-scale RCTs for some rare

cancers. Additionally, recent research revealed the increasing use of single-arm pivotal trials which could shortened drug development times.(16) Pharmaceutical companies may adopt relevant strategies to facilitate market entry, as we observed an increasing proportion of original indications supported by single-arm trials over time.

We found the judgement of RCTs to have high risk of bias or some concerns was primarily in risk of bias due to missing outcome data.(6) One potential explanation is that for placebo-controlled cancer trials, patients in the placebo arm may be more likely to withdraw consent than those in the experimental group,(26) which may result in imbalanced censoring between treatment groups and bias trial results. This issue is more common in RCTs with surrogate endpoints, some of which contain imaging or laboratory progression that cannot be assessed if the patient withdraws. We also found that bias in the randomization process was associated with exaggerated effect sizes in surrogate endpoint RCTs, (18, 27) but not in those using OS endpoints, a more objective and the gold standard cancer drug efficacy measure. Additionally, we observed a high prevalence of "some concerns" ratings in the deviations from intended interventions domain, where concerns often arose from insufficient blinding.(17) For the domain of selection of reported results, the main concerns stemmed from the unavailability of protocols for one-third of the trials and inconsistencies between reported outcomes and those specified in the protocols.

Our subgroup analysis showed a similar proportion of trials assessed as having some concern or high risk of bias among cancer drugs authorized by the FDA/EMA and those approved in China only. Most cancer drugs approved by NMPA only were supported by regional multicenter trials, in which the proportion of RCTs with some concern or high risk of bias was lower than in international multicenter trials. This could be interpreted in the context of significant system-wide efforts to improve the implementation of high-quality clinical trials since 2015.(28) Although regional multicenter trials, compared to international multicenter ones, are more likely to be implemented, the failure of approval in sintilumab (for non-small cell lung cancer, pivotal trials recruited in China only) by the FDA showed that international multicenter trials might be necessary for approvals by

other jurisdictions.(29) For novel drugs pursuing market authorization in other countries, China's domestic pharmaceutical companies should pay more attention to the design and conduct of relevant trials to avoid the risk of bias and ensure unbiased efficacy estimation when performing international multicenter studies.

Our findings highlight the need to improve the design, conduct, analysis, and reporting of pre-approval trials supporting new cancer drug approvals in China. While recent regulatory reform has narrowed the drug approval lag time with major regulatory agencies and increased the number of approvals,(25) this progress has often relied on single-arm trials and surrogate endpoints, suggesting that this acceleration may be achieved at the cost of evidence certainty. For pivotal single-arm trials where the inability to control bias is the major concern,(15) we found that external controls were frequently used without clear justification or adjustment for confounding in the regulatory review documents, limiting the credibility of effect estimates. China's regulators should make their assessments of control arm quality more specific and transparent. For RCTs, the Cochrane RoB 2 tool can serve as a structured framework to guide both sponsors and regulators in identifying critical design and implementation flaws. When assessing trial quality, particular attention should be paid to the randomization process, deviations from intended interventions, and outcome measurement. Early regulatory engagement, such as enhanced scientific advice or protocol consultation during the investigational new drug phase, could help prevent avoidable methodological issues. International collaboration may also improve evidence standards. For instance, the FDA's Project Orbis facilitated patient access to new cancer drugs through a concurrent review program of innovative cancer therapies across multiple countries.(30) While confidentiality barriers may limit full participation,(31) China's regulator could benefit from joining a similar initiative and referring to other regulators' review findings.

This study has several limitations. First, pivotal studies were identified based on regulatory review documents and product labels, which were incomplete or unclear for some indications, potentially introducing misclassification. Second, the risk-of-bias assessments relied on RoB 2.0, which, while

widely accepted, may not fully capture the methodological adequacy of oncology trials in regulatory contexts. Third, despite independent assessments of trial risk of bias by two investigators, subjectivity in certain RoB 2 aspects remains. Limited protocol availability as article supplements also meant that some risk of bias judgments were based on potential outdated versions from trial registers or pharmaceutical company websites. Fourth, the regression results on the association between risk of bias and treatment effects were exploratory, as the sample size was small and heterogeneity between trials existed. Moreover, potential correlations between risk-of-bias domains limit the interpretability of domain-specific analyses, which are not intended for causal inference. Fifth, the study covered approvals from 2017 to 2021 to allow sufficient follow-up for data availability. As more flexible trial designs and regulatory reforms are increasingly used for new indications, a timely evaluation of pivotal trials supporting recent approvals will be needed in future studies.

Acknowledgment

This study was funded by the National Natural Science Foundation of China (grant number 72274004). Dr Joseph S Ross receives research funding from the U.S. National Institutes of Health, unrelated to the current study. We thank Fengqi Liu and Dr Feng Sun from the Department of Epidemiology and Biostatistics, School of Public Health, Peking University for providing methodology consulting on risk of bias evaluation in this paper.

Article Information

Author contributions

XG and YZ conceptualized and designed the study. YZ and DC collected data and carried out the initial analyses. YZ drafted the initial manuscript. XG, MF, HN, AKW, and JSR reviewed and revised the manuscript. LS supervised this study. YZ and DC directly accessed and verified the underlying data reported in the manuscript. All authors approved the final manuscript as submitted and agreed to be accountable for all aspects of the work.

Declaration of interest

MF has received research support from the National Natural Science Foundation of China (grant number 72304011), as has LS (grant number 82273899), both outside of the current study. HN reports grants from the National Institute for Health and Care Research, European Commission, and the Commonwealth Fund outside the submitted work, and consulting fees from the World Health Organization, Arnold Ventures, and *The BMJ* (serving as an adviser). AKW reports grants from the American Cancer Society, Arnold Ventures, and the CDC and payment or honoraria from Point32Health, Tampere University, and the Finnish Medical Association. JSR currently receives research support through Yale University from Johnson and Johnson to develop methods of clinical trial data sharing, from the Food and Drug Administration for the Yale-Mayo Clinic Center for Excellence in Regulatory Science and Innovation (CERSI) program (U01FD005938), from the Agency for Healthcare Research and Quality (R01HS022882), and from Arnold Ventures; formerly received research support from the Medical Device Innovation Consortium as part of the National Evaluation System for Health Technology (NEST); and in addition, JSR was an expert witness at the

request of Relator's attorneys, the Greene Law Firm, in a qui tam suit alleging violations of the False Claims Act and Anti-Kickback Statute against Biogen Inc. that was settled September 2022. The other authors declare no competing interests.

Data sharing

Detailed data describing the sample characteristics and justifications for risk of bias assessments are available within the manuscript and appendix. All other deidentified extracted data and analytic code will be made available upon reasonable request to the corresponding author for research purposes.

References

1. National Medical Products Administration Center for Drug Evaluation. Technical guidance for cancer drug clinical trials. 2012.
2. Tannock IF, Amir E, Booth CM, Niraula S, Ocana A, Seruga B, et al. Relevance of randomised controlled trials in oncology. Lancet Oncol. 2016;17(12):e560-e7.
3. U.S. Food and Drug Administration. Guidance for Industry -Demonstrating Substantial Evidence of Effectiveness With One Adequate and Well-Controlled Clinical Investigation and Confirmatory Evidence (Draft). 2023.
4. Brown BL, Mitra-Majumdar M, Joyce K, Ross M, Pham C, Darrow JJ, et al. Trends in the Quality of Evidence Supporting FDA Drug Approvals: Results from a Literature Review. J Health Polit Policy Law. 2022;47(6):649-72.
5. Gloy V, Schmitt AM, Düblin P, Hirt J, Axfors C, Kuk H, et al. The evidence base of US Food and Drug Administration approvals of novel cancer therapies from 2000 to 2020. Int J Cancer. 2023;152(12):2474-84.
6. Naci H, Davis C, Savović J, Higgins JPT, Sterne JAC, Gyawali B, et al. 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. Bmj. 2019;366:l5221.
7. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane handbook for systematic reviews of interventions. Second;2;2nd; ed. Hoboken, NJ: Wiley-Blackwell; 2019.
8. Xu L, Gao H, Kaitin KI, Shao L. Reforming China's drug regulatory system. Nat Rev Drug Discov. 2018;17(12):858-9.
9. Senn S. Seven myths of randomisation in clinical trials. Stat Med. 2013;32(9):1439-50.
10. Zhang Y, Naci H, Wagner AK, Xu Z, Yang Y, Zhu J, et al. Overall Survival Benefits of Cancer Drugs Approved in China From 2005 to 2020. JAMA Netw Open. 2022;5(8):e2225973.
11. Yang J, Yang J, Hu YJ. Characteristics of clinical trials of new oncology drugs approved in China. Cancer. 2024;130(5):671-82.
12. National Drug Code Data File [Available from: https://www.nmpa.gov.cn/zwfw/zwfwzxfw/zxfwsjxz/index.html.
13. National Medical Products Administration Center for Drug Evaluation. Information Disclosure [Available from: https://www.cde.org.cn/.
14. Jahanshahi M, Gregg K, Davis G, Ndu A, Miller V, Vockley J, et al. The Use of External Controls in FDA Regulatory Decision Making. Ther Innov Regul Sci. 2021;55(5):1019-35.
15. The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. Guidance for Industry- E10 Choice of Control Group in Clinical Trials 2000 [Available from: https://database.ich.org/sites/default/files/E10_Guideline.pdf.
16. National Medical Products Administration Center for Drug Evaluation. Technical Guidance on the Applicability of Single-Arm Clinical Trial Designs for Cancer Drug Approval. 2023.
17. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. Bmj. 2019;366:l4898.

18. Savović J, Jones HE, Altman DG, Harris RJ, Jüni P, Pildal J, et al. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. Ann Intern Med. 2012;157(6):429-38.
19. Savovic J, Turner RM, Mawdsley D, Jones HE, Beynon R, Higgins JPT, et al. Association Between Risk-of-Bias Assessments and Results of Randomized Trials in Cochrane Reviews: The ROBES Meta-Epidemiologic Study. Am J Epidemiol. 2018;187(5):1113-22.
20. Toews I, Anglemyer A, Nyirenda JL, Alsaid D, Balduzzi S, Grummich K, et al. Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials: a meta-epidemiological study. Cochrane Database Syst Rev. 2024;1(1):Mr000034.
21. Siersma V, Als-Nielsen B, Chen W, Hilden J, Gluud LL, Gluud C. Multivariable modelling for meta-epidemiological assessment of the association between trial quality and treatment effects estimated in randomized clinical trials. Stat Med. 2007;26(14):2745-58.
22. Viechtbauer W. Conducting meta-analyses in R with the metafor package. Journal of statistical software. 2010;36:1-48.
23. Michaeli DT, Michaeli T, Albers S, Michaeli JC. Clinical trial design and treatment effects: a meta-analysis of randomised controlled and single-arm trials supporting 437 FDA approvals of cancer drugs and indications. BMJ Evid Based Med. 2024;29(5):333-41.
24. 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(1):169-77.
25. Wei Y, Zhang Y, Xu Z, Wang G, Zhou Y, Li H, et al. Cancer drug indication approvals in China and the United States: a comparison of approval times and clinical benefit, 2001-2020. Lancet Reg Health West Pac. 2024;45:101055.
26. Hillman SL, Jatoi A, Strand CA, Perlmutter J, George S, Mandrekar SJ. Rates of and Factors Associated With Patient Withdrawal of Consent in Cancer Clinical Trials. JAMA Oncol. 2023;9(8):1041-7.
27. Wood L, Egger M, Gluud LL, Schulz KF, Jüni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. Bmj. 2008;336(7644):601-5.
28. Huang H, Hou Y, Fang H, Xu L, Yu Y, Zhang H, et al. Unveiling quality of clinical trial in China: from concern to confirmation. Cancer Commun (Lond). 2024;44(5):576-9.
29. Benjamin DJ, Prasad V, Lythgoe MP. FDA decisions on new oncological drugs. Lancet Oncol. 2022;23(5):585-6.
30. de Claro RA, Spillman D, Hotaki LT, Shum M, Mouawad LS, Santos GML, et al. Project Orbis: Global Collaborative Review Program. Clin Cancer Res. 2020;26(24):6412-6.
31. Russell Flannery. U.S., China Advance Discussions On Pact To Accelerate Cancer Drug Trials 2022 [Available from: https://www.forbes.com/sites/russellflannery/2022/12/17/us-china-advance-discussions-on-pact-to-accelerate-cancer-drug-trials/.

Figures

Figure 1 title: Flowchart of sample cancer drug, indication, and pivotal trial identification

Figure 1 legend:

Abbreviation: *NMPA, National Medical Products Administration.*

Figure 2 title: Number of pivotal efficacy studies supporting new cancer drug indication approvals by National Medical Products Administration, 2017-2021

Figure 2 legend:

A. Number of cancer indications supported by randomized controlled trials or single-arm trials.

B. Number of cancer indications supported by randomized controlled trials at low risk of bias.

Note: *For randomized clinical trials, the risk of bias assessments was based on the primary efficacy endpoints.*

Figure 3 title: Risk of bias of pivotal randomized clinical trials and ability to minimize bias of pivotal single-arm trials supporting cancer indication approval by National Medical Products Administration, 2017-2021.

Figure 3 legend:

A. Risk of bias of pivotal randomized controlled trials by domain.

B. Use of external control and adjustment of confounders in pivotal single-arm trials.

Note:

[a.] *Risk of bias assessments were based on the primary efficacy endpoints.*

[b.] *P-value for the Chi-square test and Fisher's exact test.*

Figure 4 title: Comparison of trial treatment effects among different domains of risk of bias of (A) all pivotal randomized controlled trials with binary outcome, (B) with overall survival, and (C) with surrogate endpoint as the primary endpoint

Figure 4 legend:

Abbreviation: *CI, confidence interval.*

Note:

[a.] *a ratio of hazard ratio <1 suggests a larger treatment effect in trials with some concern or high risk of bias because progression or death events in oncology trials are unfavorable.*

[b.] *multivariable meta-regression model adjusting for different domains of risk of bias, cancer site, trial endpoint, trial comparator, and trial location. (for Fig4A model: R^2= 78·58%, I^2= 59·45%; for Fig4B model: R^2= 99·98%, I^2= 0·01%; for Fig4C model: R^2=69·67%, I^2=65·73%).*