Assessing the Justification, Funding, Success, and Survival Outcomes of Randomized Noninferiority Trials of Cancer Drugs
A Systematic Review and Pooled Analysis

Bishal Gyawali, MD, PhD; Frazer A. Tessema, BA; Emily H. Jung, AB; Aaron S. Kesselheim, MD, JD, MPH

Abstract

IMPORTANCE Noninferiority trials test whether a new intervention is not worse than the comparator by a given margin.

OBJECTIVES To study the characteristics of published randomized noninferiority trials in oncology with overall survival as an end point, to assess the association of justification and success in achieving noninferiority with the funding of these trials, and to evaluate the association of such trials with patient survival.

DATA SOURCES A systematic search of PubMed and Google Scholar databases was conducted in March 2018, with no date restrictions.

STUDY SELECTION Randomized noninferiority trials of cancer drug therapies with overall survival as an end point were included. Trials of decision support, supportive care, and nondrug treatment in both arms were excluded.

DATA EXTRACTION AND SYNTHESIS This study followed the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines for meta-epidemiological studies. Studies were screened for eligibility criteria, and data on criteria for noninferiority, funding, success (achieving noninferiority), and hazard ratios with confidence intervals for overall survival were extracted. Hazard ratios for overall survival were pooled across trials using a random-effects model.

MAIN OUTCOMES AND MEASURES Associations of the justification for using a noninferiority design and success in achieving noninferiority with the source of funding were assessed. Overall pooled hazard ratios and confidence intervals for overall survival were calculated.

RESULTS Among 74 noninferiority trials of cancer drug therapies, 23 (31%; enrolling 21,437 patients) used overall survival as the primary end point. The noninferiority margins for the hazard ratio of overall survival ranged from 1.08 to 1.33. Noninferiority design was justified in 14 trials (61%) but not in 9 (39%). Overall, 18 trials (78%) concluded with a finding of noninferiority. Industry funding was associated with lack of justification but not with success in achieving noninferiority. When the hazard ratios across the trials were pooled, there was no beneficial or detrimental association with overall survival, with a pooled hazard ratio of 0.97 (95% CI, 0.92-1.02).

CONCLUSIONS AND RELEVANCE The findings suggest that a substantial fraction of noninferiority trials in oncology, most of which are industry funded, lack justification for such a design. Greater

Key Points

Question Is the source of funding associated with the justification for using noninferiority design and success in claiming noninferiority in cancer drug trials, and are these trials associated with changes in patient survival?

Findings In this systematic review and pooled analysis of 23 randomized noninferiority trials of cancer drugs, which used overall survival as the end point and enrolled 21,437 patients, industry funding was associated with lack of justification but not with success in achieving noninferiority. No association of noninferiority trials with patient survival was found.

Meaning Greater regulatory attention should be paid to randomized noninferiority trials of cancer drugs, especially regarding the justification for using such a design.
Abstract (continued)

attention to the use of noninferiority designs in randomized clinical trials of cancer drugs from local and national regulators is warranted.

JAMA Network Open. 2019;2(8):e199570. doi:10.1001/jamanetworkopen.2019.9570

Introduction

Randomized clinical trials designed with a noninferiority hypothesis test whether the experimental treatment is worse than the standard of care by a given margin called the noninferiority margin.\(^1\) Noninferiority trials are useful to test interventions that aim at providing compensatory benefits without necessarily compromising efficacy. If the new drug or strategy reduces cost or adverse effects or improves the ease of administration, patients may reasonably accept the possibility of slight reductions in efficacy.

However, there are numerous concerns associated with defining, designing, conducting, reporting, and interpreting noninferiority trials,\(^1,2\) including in oncology.\(^3\) For example, how much of a possible reduction in efficacy is reasonable? How noninferiority limits are defined may raise ethical issues related to the sufficiency of patient informed consent.\(^4\) In some cases, noninferiority designs have been useful. For example, the standard of care for the use of zoledronic acid for preventing skeletal-related events in patients with cancer and bone metastases used to be an injection every month. Noninferiority trials have now established that the frequency could safely be reduced to once every 3 months without compromising efficacy outcomes.\(^5,6\)

Noninferiority trials have been applied sparingly to oncology drugs because patients with cancer are unlikely to accept even the possibility of reduced efficacy of their oncology drugs. However, in a 2018 case, a noninferiority trial served as the pivotal trial for US Food and Drug Administration approval of a new cancer drug, even though the new drug did not necessarily provide any benefits in terms of cost, ease of administration, or toxic effects.\(^7\) Previous studies have shown that industry-funded trials were more likely to use a noninferiority design than nonindustry-sponsored trials.\(^8\)

In cancer treatment, noninferiority trials with overall survival (OS) outcomes are most critical from a clinical and regulatory point of view because it would be unwise to use a new therapy based on its noninferiority regarding response rates or progression-free survival without knowing the effects on OS, as the drug might very well have more substantially inferior outcomes on survival. Similarly, patients are most concerned with compromise in survival when the results from noninferiority trials are translated in clinics.

To understand the characteristics of noninferiority trials used in evaluating cancer drugs, we sought to systematically review the use of the noninferiority hypothesis in randomized clinical trials in the field of oncology. In this systematic review and pooled analysis, we investigated the basis or reasoning for using noninferiority designs, the funding of these trials, and efficacy outcomes.

Methods

The objectives of this study were as follows: (1) to study the characteristics of randomized cancer drug trials that use noninferiority design, including the success rate and justification for using noninferiority designs, (2) to study whether the success in claiming noninferiority or the lack of justification for using noninferiority design was associated with the source of funding, and (3) to study the overall association of the drug being tested with patients’ survival. This study was conducted in accordance with the modification of Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) reporting guideline for meta-epidemiological studies.\(^9\)
Study Identification
We conducted a systematic search of the PubMed database in March 2018 without date restrictions, supplemented with a search of the Google Scholar database. To identify published trials in cancer that used a noninferiority design, we used the search terms neoplasms or cancer or tumor* or malignan* or oncology AND non inferior* or non-inferior* or noninferior* and limited our search to findings published in English. After title and abstract screening by 2 of us (B.G. and F.A.T.) acting independently, the full texts of potentially relevant studies were downloaded and reviewed for the following exclusion criteria: (1) not a randomized design; (2) trials in pediatric populations; (3) trials of surgery or radiotherapy only in all treatment arms (trials comparing a drug in 1 arm and surgery or radiotherapy in other arms were eligible); (4) trials of decision support, diagnostic modalities, or supportive care; (5) trials assessing behavioral interventions, genetic counseling, screening, or diagnostic modalities; (6) trials assessing only pharmacokinetics; and (7) post hoc analyses. For this study, we limited our analysis to trials with a primary or a coprimary end point of OS for 2 reasons. First, the most important concern for patients making treatment decisions based on noninferiority trials is the potential for compromise in survival, rather than compromise in surrogate measures. Second, we planned to conduct a pooled analysis, and it would not be possible to pool survival with other surrogate measures across the trials.

Data Extraction
Data were independently extracted from published reports by 2 of us (F.A.T. and E.H.J) and verified by a third author (B.G.), with discrepancies resolved through consensus of all authors. We collected year of publication, treatment setting, primary end point, sample size, blinding (ie, double blind or open label), funding (ie, public, industry, or mixed public and industry), and the authors’ listed criteria for using a noninferiority design.

We judged whether the noninferiority design was justified for each trial. Noninferiority was considered justified if the intervention being tested provided at least 1 of the following benefits to patients: (1) less cost; (2) decreased frequency of administration; (3) increased ease of administration, such as noninjectable (eg, oral) vs injectable formulation; or (4) improved quality of life.

We also extracted information on the outcome of the trial in terms of whether noninferiority was achieved. A trial was considered successful if it achieved noninferiority based on its own criteria. For trials that achieved noninferiority, we checked if the intervention also proved superiority based on 95% CIs and whether the publication concluded superiority. For quality-of-life outcomes, we considered quality of life to be improved if the summary was statistically better; we did not examine each domain of the assessment tool separately. Finally, information on the hazard ratio (HR) and 95% CI for OS were extracted from the published reports for pooled analysis. If a CI different from a 95% CI was reported, we recalculated the CI to a 95% CI.

Statistical Analysis
The associations of the justification for using the noninferiority design and success in achieving noninferiority with the funding source were assessed using Fisher exact tests. The overall association of the trial drugs on OS was assessed by pooling the HRs across the trials using a random-effects meta-analysis to account for heterogeneity. Heterogeneity among studies was assessed using the Cochrane Q statistic (assumption of homogeneity was considered invalid for values of P < .10) and quantified using an I² test. Subgroup analyses were prespecified and included funding, blinding, and success. All statistical analyses were conducted using Stata version 15 (StataCorp), and a 2-sided P < .05 was considered statistically significant.
Results

Among 128 randomized noninferiority trials in oncology in adults identified through our search, 74 (58%) were drug trials. Of those, 23 (31%) enrolled 21 437 patients, used OS as the primary or coprimary end point, and were therefore included in our analysis (Figure 1 and Table 1).10-32 None of these trials were blinded. Approximately half (12 [52%]) had industry funding, 5 (22%) were publicly funded, and 6 (26%) had mixed funding.

Noninferiority was defined in terms of the upper limit of the CI of the HR for OS in 20 (87%) trials. Most trials used 95% CIs,11-14,20,22-26,28-32 but some used 90% CIs,17,18,21 96% CIs,19 and 80% CIs10 (Table 2). A 2012 trial16 defined noninferiority in terms of difference in 5-year OS rates, a 2016 trial as retaining at least 50% of the OS effects of the control arm,27 and a 2012 trial in terms of the lower limit of the 95% CI for median OS greater than median minus 10% of the overall survival of the control arm15 (Table 2).

The upper limit of the CI for the HR for OS in the 20 trials that used that measure as their criteria for noninferiority ranged from 1.08 to 1.33. Of these, 14 trials (70%)12-14,17,20-22,24,25,28,30-32 had a limit of 1.2 or more, meaning that up to a 20% increase in mortality risk was considered noninferior (ie, acceptable). A 2017 trial10 had 3 arms comparing 2 different regimens in noninferiority design against the control.30 Quality of life was improved in 8 trials (35%),10,12,17,19,21,22,31,32 not improved in another 8 (35%),13,14,18,23-25,28,30 and not assessed or assessed but not reported in the remaining 7 (30%) (Table 1).11,15,16,20,26,27,29

Justification for Noninferiority

Overall, 14 trials (61%) met our criteria for justifying the noninferiority design (Table 1).10-14,16-19,21,22,25,33,32 A new oral drug vs injectable standard was the most common justification (7 trials)11,16,19,22,25,31,32 for using the noninferiority design, followed by intermittent vs continuous therapies (4 trials).10,12,13,17 Other justifications included avoiding surgery with chemotherapy18 or using carboplatin instead of cisplatin.14,21

Figure 1. Flow Diagram for the Selection of Studies

569 Records identified through database searching

14 Additional records identified through other sources

121 Duplicate records removed

462 Records after duplicates removed

100 Records excluded from title screening

362 Abstracts screened

234 Records excluded after abstract screening

128 Full-text articles assessed for eligibility

54 Full-text articles excluded

74 Studies of cancer drugs that met eligibility criteria

51 Studies excluded for not having overall survival endpoint

23 Studies included in analysis
Among the 9 noninferiority trials (39%) that did not meet our justification criteria, comparisons were made between oral-oral or intravenous-intravenous drugs, usually of the same class. Most trials were industry funded, while 1 trial had mixed funding. All publicly funded noninferiority trials met the criteria for justification. Lack of justification was associated with funding source (P = .02, assessed using the Fisher exact test).

**Success in Establishing Noninferiority**
A total of 18 trials (78%) established noninferiority for the intervention drug (Table 2). One trial proved superiority. Of the 12 industry-funded trials, 10

### Table 1. Justifications for Noninferiority Trials of Cancer Drugs in Oncology With Overall Survival Endpoint

| Source | Total Participants, No. | Cancer Type | Patients, No. | QoL Results | Justification of Noninferiority Design |
|--------|-------------------------|-------------|---------------|-------------|---------------------------------------|
| Adams et al, 2011 | 2445 | Colorectal | 815 | 815 | Better | Justified: intermittent vs continuous oxaliplatin |
| Boku et al, 2009 | 468 | Gastric | 234 | 234 | Not assessed | Justified: oral S-1 vs IV 5-FU; S-1 less expensive |
| Crook et al, 2012 | 1386 | Prostate | 690 | 696 | Better | Justified: intermittent vs continuous androgen deprivation therapy |
| daSilva et al, 2014 | 918 | Prostate | 462 | 456 | Not statistically different | Justified: intermittent vs continuous androgen deprivation therapy |
| Ferry et al, 2017 | 909 | NSCLC | 453 | 456 | Not different; statistical comparison not provided | Justified: carboplatin vs cisplatin |
| Fink et al, 2012 | 795 | SCLC | 346 | 334 | Not assessed | Not justified: topotecan-cisplatin vs cisplatin-etoposide |
| Hofheinz et al, 2012 | 392 | Colorectal | 195 | 197 | Not assessed | Justified: oral capecitabine vs IV 5-FU |
| Hussain et al, 2013 | 1535 | Prostate | 770 | 765 | Slightly better but summary statistic not available | Justified: intermittent vs continuous hormone therapy |
| Kehoe et al, 2015 | 500 | Ovarian | 274 | 276 | Not statistically different | Justified: avoiding surgery |
| Kim et al, 2008 | 1433 | NSCLC | 723 | 710 | Better | Justified: oral gefitinib vs IV docetaxel |
| Kim et al, 2009 | 491 | Colorectal | 246 | 245 | Not reported | Not justified: FOLFOX vs irinotecan |
| Kitagawa et al, 2015 | 253 | Cervical | 126 | 127 | Better | Justified: carboplatin vs cisplatin; less hospitalization time |
| Kubota et al, 2015 | 608 | NSCLC | 301 | 295 | Better | Justified: oral S-1 vs IV docetaxel in combination with platinum |
| Kudo et al, 2018 | 1492 | Liver | 476 | 475 | Not statistically different | Not justified: lenvatinib vs sorafenib; lenvatinib costs more than sorafenib and offers no other benefits |
| Lang et al, 2013 | 564 | Breast | 279 | 285 | Not statistically different | Not justified: bevacizumab plus capcitabine vs bevacizumab plus paclitaxel; capcitabine and paclitaxel are both established as first-line agents; the control arm of bevacizumab plus paclitaxel is now withdrawn by the FDA |
| Okamoto et al, 2010 | 564 | NSCLC | 282 | 281 | Not statistically different | Justified: oral S-1 vs IV paclitaxel in combination with carboplatin |
| Popov et al, 2008 | 1921 | Colorectal | 952 | 969 | Not assessed | Not justified: IV ralitrexed vs IV 5-FU |
| Price et al, 2016 | 1010 | Colorectal | 596 | 594 | Not assessed | Not justified: panitumumab vs cetuximab |
| Satouchi et al, 2014 | 284 | SCLC | 142 | 142 | Not statistically different | Not justified: amrubin-platinum vs irinotecan-platinum |
| Scaglotti et al, 2008 | 1725 | NSCLC | 862 | 863 | Not assessed | Not justified: cisplatin-pemetrexed vs cisplatin-gemcitabine; pemetrexed offered a better profile for some adverse effects but was much more expensive than gemcitabine |
| Shitara et al, 2017* | 741 | Gastric | 246 and 247 | 248 | Statistical comparison not provided | Not justified: abraxane 3 times per week vs abraxane weekly vs paclitaxel weekly; abraxane costs more than 50-fold what paclitaxel costs; premedication not necessary with abraxane |
| Takashima et al, 2016 | 618 | Breast | 309 | 309 | Better | Justified: oral S-1 vs IV taxane |
| Uesaka et al, 2016 | 385 | Pancreas | 192 | 193 | Better | Justified: oral S-1 vs IV gemcitabine |

Abbreviations: FDA, US Food and Drug Administration; FOLFOX, 5-fluorouracil plus leucovorin plus oxaliplatin; 5-FU, 5-fluorouracil; IV, intravenous; NSCLC, non-small cell lung cancer; QoL, quality of life; SCLC, small cell lung cancer.

* The trial by Shitara et al had 3 arms.
established noninferiority for the drug being tested.\textsuperscript{15,16,19,20,22,23,25,27,29,30} Success in achieving noninferiority was not associated with funding (\(P = .80\), assessed using the Fisher exact test).

**Association With OS**

The HR was more than 1 in 10 trials (41%),\textsuperscript{10,12,17,19,22,24,26,28,30,31} but was significant in only 1 trial\textsuperscript{28} (Table 2). When the HRs across trials were pooled using random-effects meta-analysis, there was no beneficial or detrimental association with patient survival (pooled HR, 0.97; 95% CI, 0.92-1.02); the heterogeneity among the trials was substantial as trials across different tumor types were pooled (\(I^2 = 53.8\%), \(P = .001\)) (Figure 2). This analysis was repeated using comparisons for each cohort of the trial by Shitara et al\textsuperscript{30} (the trial with 3 arms) independently and results did not change. On subgroup analysis, there was no difference between industry-funded trials vs mixed or publicly funded trials as well as no difference between trials in which noninferiority design was justified vs not justified.

**Discussion**

In systematically searching for noninferiority trials of cancer drugs, we found that only 31% used OS as the primary end point. Among these trials, the criteria to define noninferiority varied from 1.08 to 1.33 for the upper limit of the CI of the HR for death. Ease of administration (oral formulation vs injectable control drugs) was the most common justification for the noninferiority design, while we found insufficient justification for 40% of such trials. Most trials were successful in proving

| Table 2. Outcomes of Noninferiority Trials of Cancer Drugs in Oncology With OS End Point |
| Source | Cancer Type | Funding | Noninferiority Criteria for Hazard Ratio | Success | OS, HR (95% CI) |
|--------|-------------|---------|----------------------------------------|---------|----------------|
| Adams et al,\textsuperscript{10} 2011\textsuperscript{a} | Colorectal | Mixed | 1.162\textsuperscript{b} | No | 1.08 (1.01-1.17) |
| Boku et al,\textsuperscript{11} 2009 | Gastric | Mixed | 1.16 \textsuperscript{c} | Yes | 0.83 (0.68-1.01) |
| Crook et al,\textsuperscript{12} 2012 | Prostate | Public | 1.25 \textsuperscript{c} | Yes | 1.02 (0.86-1.21) |
| daSilva et al,\textsuperscript{13} 2014 | Prostate | Public | 1.21 \textsuperscript{c} | Yes | 0.90 (0.76-1.07) |
| Ferry et al,\textsuperscript{14} 2017 | NSCLC | Mixed | 1.2 \textsuperscript{c} | Yes | 0.93 (0.83-1.04) |
| Fink et al,\textsuperscript{15} 2012 | SCLC | Industry | Lower limit of 95% CI of median OS should be >10% less than the median OS of control arm | Yes | 0.93 (0.79-1.10) |
| Hoffheinz et al,\textsuperscript{16} 2012 | Colorectal | Industry | 5-y OS rate difference 12.5% | Yes | 0.67 (0.44-1.00) |
| Hussain et al,\textsuperscript{17} 2013\textsuperscript{a} | Prostate | Public | 1.2 \textsuperscript{d} | No | 1.10 (0.99-1.23) |
| Kehoe et al,\textsuperscript{18} 2015\textsuperscript{a} | Ovarian | Public | 1.18 \textsuperscript{d} | Yes | 0.87 (0.72-1.05) |
| Kim et al,\textsuperscript{19} 2008\textsuperscript{a} | NSCLC | Industry | 1.154 \textsuperscript{d} | Yes | 1.02 (0.91-1.15) |
| Kim et al,\textsuperscript{20} 2009 | Colorectal | Industry | 1.33 \textsuperscript{e} | Yes | 0.92 (0.80-1.10) |
| Kitagawa et al,\textsuperscript{21} 2015\textsuperscript{a} | Cervical | Public | 1.29 \textsuperscript{e} | Yes | 0.99 (0.79-1.25) |
| Kubota et al,\textsuperscript{22} 2015 | NSCLC | Industry | 1.322 \textsuperscript{e} | Yes | 1.01 (0.84-1.23) |
| Kudo et al,\textsuperscript{23} 2018 | Liver | Industry | 1.08 \textsuperscript{f} | Yes | 0.92 (0.79-1.06) |
| Lang et al,\textsuperscript{24} 2013 | Breast | Mixed | 1.33 \textsuperscript{e} | No | 1.06 (0.80-1.40) |
| Okamoto et al,\textsuperscript{25} 2010\textsuperscript{a} | NSCLC | Industry | 1.33 \textsuperscript{f} | Yes | 0.93 (0.69-1.17) |
| Popov et al,\textsuperscript{26} 2008 | Colorectal | Industry | 1.25 \textsuperscript{e} | No | 1.01 (0.84-1.23) |
| Price et al,\textsuperscript{27} 2016 | Colorectal | Industry | ±50% Retention of OS effects of control | Yes | 0.94 (0.82-1.07) |
| Satouchi et al,\textsuperscript{28} 2014 | SCLC | Industry | 1.31 \textsuperscript{e} | No | 1.43 (1.10-1.85) |
| Scagliotti et al,\textsuperscript{29} 2008 | NSCLC | Industry | 1.176 \textsuperscript{e} | Yes | 0.94 (0.84-1.05) |
| Shitara et al,\textsuperscript{30} 2017\textsuperscript{f} | Gastric | Industry | 1.25 \textsuperscript{e} | Yes for arm A; no for arm B | Arm A, 0.97 (0.76-1.23); arm B, 1.06 (0.87-1.31) |
| Takashima et al,\textsuperscript{31} 2016 | Breast | Mixed | 1.333 \textsuperscript{e} | Yes | 1.05 (0.86-1.27) |
| Uesaka et al,\textsuperscript{32} 2016 | Pancreas | Mixed | 1.25 \textsuperscript{e} | Yes | 0.57 (0.44-0.72) |

Abbreviations: HR, hazard ratio; NSCLC, non–small cell lung cancer; OS, overall survival; SCLC, small cell lung cancer.

\(*\) Reported a CI different from a 95% CI, which was recalculated in the Table to 95% CIs.

\(\textsuperscript{b}\) Noninferiority criteria defined by upper limit of 80% CI for HR.

\(\textsuperscript{c}\) Noninferiority criteria defined by upper limit of 95% CI for HR.

\(\textsuperscript{d}\) Noninferiority criteria defined by upper limit of 90% CI for HR.

\(\textsuperscript{e}\) Noninferiority criteria defined by upper limit of 96% CI for HR.

\(\textsuperscript{f}\) The trial by Shitara et al\textsuperscript{30} had 3 arms.
noninferiority, and when the hazard ratio was pooled across these trials, there was no detrimental effect on OS.

Noninferiority trials in oncology have previously been criticized for “critical deficiencies in design and reporting.”\(^{33}\) In this cohort, OS was the primary end point in fewer than one-third of noninferiority trials. However, using a noninferiority hypothesis implies a willingness to compromise efficacy outcomes to achieve benefits in other areas. Surrogate measures, such as progression-free survival, have been shown to overestimate (or, in the case of immunotherapy drugs, underestimate)\(^{34}\) benefit and translate to smaller than expected gains in OS. In most cases these surrogate measures have been shown to lack correlation with OS.\(^{35}\) Thus, a trial testing noninferiority in a surrogate measure is ethically challenging if it is impossible to estimate the magnitude of potential compromise in efficacy and communicate that clearly to patients participating in such trials as part of the informed consent process. Investigators should therefore be wary of conducting trials assessing noninferiority on surrogate measures.

The criteria to define noninferiority among the trials in the cohort varied from 1.08 to 1.33 for the upper bound of the CI of the HR for OS, which means from an 8% to a 33% increase in the hazard of death was considered acceptable (noninferior) in these trials. Furthermore, in multiple cases, this upper limit was defined not for a 95% CI but for a 90% or even an 80% CI. As previous studies have shown, there are no set methods for determining limits defining noninferiority.\(^{3,33}\) We also found none of the 4 key justifications—lower toxic effects, lower cost, ease of administration, or better quality of life—for using noninferiority designs in approximately 40% of the cohort. Such benefits should be the primary rationale for a patient consenting to participate in a trial testing if a new treatment is not worse than the standard treatment by a prespecified margin.

Our final objective was to assess if patients enrolled in the experimental arm of noninferiority trials experienced reduced OS compared with those in the control arms. Reassuringly, we found no
such association. However, this pooled result should be interpreted with caution because, when examining individual trials, we found 10 trials (43%) in which the HR was more than 1, including 1 in which the HR was significantly more than 1. Similarly, 1 noninferiority trial actually proved superiority for the experimental arm. Subgroup analyses revealed no differences in OS based on funding or reporting 1 of the 4 key justifications for noninferiority.

Half of the trials in our study had industry funding, and having industry funding was significantly associated with missing the justifications we identified for the noninferiority design. Industry-funded noninferiority trials also were successful in proving noninferiority in 83% of cases, although the association of success with funding was not significant. Our results are similar to those of a previous analysis showing that 43% of noninferiority trials in oncology were industry funded, 73% reported positive results, and OS was the primary end point in 25%. Another study that included all noninferiority trials across disciplines showed that 83% produced favorable results irrespective of funding source, similar to our findings. These results should be interpreted in light of the fact that success in achieving noninferiority depends not only on the intervention but also on the criteria used to define noninferiority, which is often arbitrary. Furthermore, we cannot rule out publication bias against noninferiority trials that failed to achieve noninferiority.

These findings are important for those helping design and oversee the conduct of noninferiority trials in oncology. Noninferiority trials may be attractive because of the high probability of success. Indeed, the noninferiority design has been described as having a low risk of failure. However, our data show that institutional review boards and drug regulators should take an active role in adjudicating whether the noninferiority design is acceptable for the given question. When noninferiority design trials are considered important, the criteria to define noninferiority should be clearly defined based on a widely accepted rationale and should incorporate patient input.

The Consolidated Standards of Reporting Trials (CONSORT) statement on reporting noninferiority trials recommends providing a rationale for the noninferiority design and the criteria for defining noninferiority; however, further research is needed to assess if this recommendation has improved reporting practices in oncology. Similarly, the US Food and Drug Administration has issued a guidance for industry on noninferiority trials. However, the US Food and Drug Administration recommends a noninferiority design be chosen “when it would not be ethical to use a placebo.” While that is necessary, it is not sufficient, and superiority design trials against an active comparator should be encouraged unless the noninferiority design is justified for other compelling reasons, such as the ones we have mentioned. The guidance could also be improved by highlighting the need to incorporate patient input on the acceptable margin for defining noninferiority among various tumor types.

Limitations
This study has limitations. Although our analysis included trials involving more than 21,000 participants, for the analysis of the association of different parameters with funding, this study was limited by the relatively small total number of trials (23). Another important limitation is our consensus-based definition for adjudicating whether the noninferiority design was justified. However, other authors have also concluded that “any new treatment, even if noninferior to standard treatment, should have some benefits, such as for quality of life, cost, or safety.” We focused on noninferiority trials testing OS end points; however, many trials also test noninferiority in surrogate measures, such as response rates. The association of such trials with patient treatment outcomes is a topic for future research.

Conclusions
Noninferiority randomized trials in oncology should be used only when there are important potential benefits that the experimental drug can offer patients. However, among 23 such trials testing OS, we found that a substantial fraction did not offer any of the 4 key criteria for justification. Greater
attention to the use of noninferiority designs in cancer drug clinical trials from local and national regulators is warranted.

ARTICLE INFORMATION
Accepted for Publication: June 30, 2019.
Published: August 30, 2019. doi:10.1001/jamanetworkopen.2019.9570
Open Access: This is an open access article distributed under the terms of the CC-BY License. © 2019 Gyawali B et al.
JAMA Network Open.

Corresponding Author: Bishal Gyawali, MD, PhD, Cancer Research Institute, Cancer Care and Epidemiology, 10 Stuart St, Level 2, Queen's University, Kingston, ON K7L 3N6, Canada (bg.bishalgyawali@gmail.com).

Author Affiliations: Division of Cancer Care and Epidemiology, Departments of Medical Oncology and Public Health Sciences, Queen's University, Kingston, Ontario, Canada (Gyawali); Program on Regulation, Therapeutics, and Law (PORTAL), Division of Pharmacoepidemiology and Pharmacoeconomics, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts (Gyawali, Tessema, Jung, Kesselheim).

Author Contributions: Dr Gyawali had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Gyawali, Kesselheim.

Acquisition, analysis, or interpretation of data: All authors.

Drafting of the manuscript: Gyawali, Tessema.

Critical revision of the manuscript for important intellectual content: Gyawali, Jung, Kesselheim.

Statistical analysis: Gyawali, Tessema.

Obtained funding: Kesselheim.

Administrative, technical, or material support: Gyawali, Tessema, Jung.

Supervision: Gyawali.

Conflict of Interest Disclosures: Dr Kesselheim reported receiving unrelated research funding from the US Food and Drug Administration Division of Health Communication (2013-2016). No other disclosures were reported.

Funding/Support: Work on this project was funded by Arnold Ventures. Dr Kesselheim is supported by the Harvard-MIT Center for Regulatory Science and the Engelberg Foundation.

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

REFERENCES
1. Mauri L, D'Agostino RB Sr. Challenges in the design and interpretation of noninferiority trials. N Engl J Med. 2017;377(14):1357-1367. doi:10.1056/NEJMra1510063

2. Rehal S, Morris TP, Fielding K, Carpenter JR, Phillips PP. Non-inferiority trials: are they inferior? a systematic review of reporting in major medical journals. BMJ Open. 2016;6(10):e012594. doi:10.1136/bmjopen-2016-012594

3. Burotto M, Prasad V, Fojo T. Non-inferiority trials: why oncologists must remain wary. Lancet Oncol. 2015;16(4):364-366. doi:10.1016/S1470-2045(15)70129-4

4. Riechelmann RP, Alex A, Cruz L, Bariani GM, Hoff PM. Non-inferiority cancer clinical trials: scope and purposes underlying their design. Ann Oncol. 2013;24(7):1942-1947. doi:10.1093/annonc/mdt073

5. Himelstein AL, Foster JC, Khatcheressian JL, et al. Effect of longer-interval vs standard dosing of zoledronic acid on skeletal events in patients with bone metastases: a randomized clinical trial. JAMA. 2017;317(1):48-58. doi:10.1001/jama.2016.19425

6. Hortobagyi GN, Van Poznak C, Harker WG, et al. Continued treatment effect of zoledronic acid dosing every 12 vs 4 weeks in women with breast cancer metastatic to bone: the OPTIMIZE-2 randomized clinical trial. JAMA Oncol. 2017;3(7):906-912. doi:10.1001/jamaoncol.2016.6316

7. Gyawali B, Kesselheim AS. US Food and Drug Administration approval of new drugs based on noninferiority trials in oncology: a dangerous precedent? JAMA Oncol. 2019;5(5):607-608. doi:10.1001/jamaoncol.2019.0093

8. Flacco ME, Manzoli L, Boccia S, et al. Head-to-head randomized trials are mostly industry sponsored and almost always favor the industry sponsor. J Clin Epidemiol. 2015;68(7):811-820. doi:10.1016/j.jclinepi.2014.12.016
9. Murad MH, Wang Z. Guidelines for reporting meta-epidemiological methodology research. *Evid Based Med*. 2017;22(4):139-142.

10. Adams RA, Meade AM, Seymour MT, et al; MRC COIN Trial Investigators. Intermittent versus continuous oxaliplatin and fluoropyrimidine combination chemotherapy for first-line treatment of advanced colorectal cancer: results of the randomised phase 3 MRC COIN trial. *Lancet Oncol*. 2011;12(7):642-653. doi:10.1016/S1470-2045(11)70102-4

11. Boku N, Yamamoto S, Fukuda H, et al; Gastrointestinal Oncology Study Group of the Japan Clinical Oncology Group. Fluorouracil versus combination of irinotecan plus cisplatin versus S-1 in metastatic gastric cancer: a randomised phase 3 study. *Lancet Oncol*. 2009;10(11):1063-1069. doi:10.1016/S1470-2045(09)70259-1

12. Crook JM, O’Callaghan CJ, Duncan G, et al. Intermittent androgen suppression for rising PSA level after radiotherapy. *N Engl J Med*. 2012;367(10):895-903. doi:10.1056/NEJMoa120546

13. Calais da Silva F, Calais da Silva FM, Gonçalves F, et al. Locally advanced and metastatic prostate cancer treated with intermittent androgen monotherapy or maximal androgen blockade: results from a randomised phase 3 study by the South European Urooncological Group. *Eur Urol*. 2014;66(2):232-239. doi:10.1016/j.eururo.2013.03.055

14. Ferry D, Billingham L, Jarrett H, et al. Carboplatin versus two doses of cisplatin in combination with gemcitabine in the treatment of advanced non-small-cell lung cancer: results from a British Thoracic Oncology Group randomised phase III trial. *Eur J Cancer*. 2017;83:302-312.

15. Fink TH, Huber RM, Heigener DF, et al; “Aktion Bronchialkarzinom” (ABC Study Group). Topotecan/cisplatin compared with cisplatin/etoposide as first-line treatment for patients with extensive disease small-cell lung cancer: final results of a randomized phase III trial. *J Thorac Oncol*. 2012;7(9):1424-1439. doi:10.1097/JTO.0b013e3182606ede7

16. Hofheinz RD, Wenz F, Post S, et al. Chemoradiotherapy with capetitabine versus fluorouracil for locally advanced rectal cancer: a randomised, multicentre, non-inferiority, phase 3 trial. *Lancet Oncol*. 2012;13(6):579-588. doi:10.1016/S1470-2045(12)70116-X

17. Hussain M, Tangen CM, Berry DL, et al. Intermittent versus continuous androgen deprivation in prostate cancer. *N Engl J Med*. 2013;368(14):1314-1325. doi:10.1056/NEJMoa122299

18. Kehoe S, Hook J, Nankivell M, et al. Primary chemotherapy versus primary surgery for newly diagnosed advanced ovarian cancer (CHORUS): an open-label, randomised, controlled, non-inferiority trial. *Lancet*. 2015;386(9990):249-257. doi:10.1016/S0140-6736(14)62223-6

19. Kim ES, Hirsh V, Mok T, et al. Gefitinib versus docetaxel in previously treated non-small-cell lung cancer (INTEREST): a randomised phase III trial. *Lancet*. 2008;372(9652):1809-1818. doi:10.1016/S0140-6736(08)67584-4

20. Kim GP, Sargent DJ, Mahoney MR, et al. Phase III noninferiority trial comparing irinotecan with oxaliplatin, fluorouracil, and leucovorin in patients with advanced colorectal carcinoma previously treated with fluorouracil: N9841. *J Clin Oncol*. 2009;27(17):2848-2854. doi:10.1200/JCO.2008.20.4552

21. Kitagawa R, Katsumata N, Shibata T, et al. Paclitaxel plus carboplatin versus paclitaxel plus cisplatin in metastatic or recurrent cervical cancer: the open-label randomized phase III trial JCOG0505. *J Clin Oncol*. 2015;33(19):2129-2135. doi:10.1200/JCO.2014.58.4391

22. Kubota K, Sakai H, Katakami N, et al; Tokyo Cooperative Oncology Group. A randomized phase III trial of oral S-1 plus cisplatin versus docetaxel plus cisplatin in Japanese patients with advanced non-small-cell lung cancer: TCGO0701 CATS trial. *Ann Oncol*. 2015;26(7):1401-1408. doi:10.1093/annonc/mdv190

23. Kudo M, Finn RS, Qin S, et al. Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. *Lancet*. 2018;391(10126):1163-1173. doi:10.1016/S0140-6736(18)30207-1

24. Lang I, Brodowicz T, Ryvo L, et al; Central European Cooperative Oncology Group. Bevacizumab plus paclitaxel versus bevacizumab plus capetitabine as first-line treatment for HER2-negative metastatic breast cancer: interim efficacy results of the randomised, open-label, non-inferiority, phase 3 TURANDOT trial. *Lancet Oncol*. 2013;14(2):125-133. doi:10.1016/S1470-2045(12)70566-1

25. Okamoto I, Yoshikawa H, Morita S, et al. Phase III trial comparing oral S-1 plus carboplatin with paclitaxel plus carboplatin in chemotherapy-naive patients with advanced non-small-cell lung cancer: results of a west Japan oncology group study. *J Clin Oncol*. 2010;28(36):5240-5246. doi:10.1200/JCO.2010.31.0326

26. Popov I, Carrato A, Sobrero A, et al. Raltitrexed (Tomudex) versus standard leucovorin-modulated bolus 5-fluourouracil: Results from the randomised phase III Pan-European Trial in Adjuvant Colon Cancer 01 (PETACC-1). *Eur J Cancer*. 2008;44(15):2204-2211.
27. Price T, Kim TW, Li J, et al. Final results and outcomes by prior bevacizumab exposure, skin toxicity, and hypomagnesaeemia from ASPECTT: randomized phase 3 non-inferiority study of panitumumab versus cetuximab in chemorefractory wild-type KRAS exon 2 metastatic colorectal cancer. *Eur J Cancer*. 2016;68:51-59.

28. Satouchi M, Kotani Y, Shibata T, et al. Phase III study comparing amrubicin plus cisplatin with irinotecan plus cisplatin in the treatment of extensive-disease small-cell lung cancer. *JCOG 0509*. *J Clin Oncol*. 2014;32(12):1262-1268. doi:10.1200/JCO.2013.53.5153

29. Scagliotti GV, Parikh P, von Pawel J, et al. Phase III study comparing cisplatin plus gemcitabine with cisplatin plus pemetrexed in chemotherapy-naive patients with advanced-stage non-small-cell lung cancer. *J Clin Oncol*. 2008;26(21):3543-3551. doi:10.1200/JCO.2007.15.0375

30. Shitara K, Takashima A, Fujitani K, et al. Nab-paclitaxel versus solvent-based paclitaxel in patients with previously treated advanced gastric cancer (ABSOLUTE): an open-label, randomised, non-inferiority, phase 3 trial. *Lancet Gastroenterol Hepatol*. 2017;2(4):277-287. doi:10.1016/S2468-1253(16)30219-9

31. Takashima T, Mukai H, Hara F, et al; SELECT BC Study Group. Taxanes versus S-1 as the first-line chemotherapy for metastatic breast cancer (SELECT BC): an open-label, non-inferiority, randomised phase 3 trial. *Lancet Oncol*. 2016;17(1):90-98. doi:10.1016/S1470-2045(15)00411-8

32. Uesaka K, Boku N, Fukutomi A, et al; JASPAC 01 Study Group. Adjuvant chemotherapy of S-1 versus gemcitabine for resected pancreatic cancer: a phase 3, open-label, randomised, non-inferiority trial (JASPAC 01). *Lancet*. 2016;388(10041):248-257. doi:10.1016/S0140-6736(16)30583-9

33. Tanaka S, Kinjo Y, Kataoka Y, Yoshimura K, Teramukai S. Statistical issues and recommendations for noninferiority trials in oncology: a systematic review. *Clin Cancer Res*. 2012;18(7):1837-1847. doi:10.1158/1078-0432.CCR-11-1653

34. Gyawali B, Hey SP, Kesselheim AS. A comparison of response patterns for progression-free survival and overall survival following treatment for cancer with PD-1 inhibitors: a meta-analysis of correlation and differences in effect sizes. *JAMA Netw Open*. 2018;1(2):e180416. doi:10.1001/jamanetworkopen.2018.0416

35. 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.

36. Soonawala D, Dekkers OM, Vandenbroucke JP, Egger M. Noninferiority is (too) common in noninferiority trials. *J Clin Epidemiol*. 2016;71:118-120. doi:10.1016/j.jclinepi.2015.11.009

37. Piaggio G, Elbourne DR, CopcoCK SJ, Evans SJ, Altman DG; CONSORT Group. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. *JAMA*. 2012;308(24):2594-2604. doi:10.1001/jama.2012.87802

38. US Food and Drug Administration. Non-inferiority clinical trials to establish effectiveness: guidance for industry. https://www.fda.gov/media/78504/download. Accessed June 29, 2019.