UCSF
UC San Francisco Previously Published Works

Title

Assessing Patient Risk, Benefit, and Outcomes in Drug Development: A Decade of Lenvatinib Clinical Trials: A Systematic Review

Permalink

https://escholarship.org/uc/item/4vv2r8pn

Journal

Targeted Oncology, 19(2)

ISSN

1776-2596

Authors

Crotty, Patrick
Kari, Karim
Hughes, Griffin K
et al.

Publication Date

2024-03-01

DOI

10.1007/s11523-024-01040-5

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial License, available at https://creativecommons.org/licenses/by-nc/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Targeted Oncology (2024) 19:161–173
https://doi.org/10.1007/s11523-024-01040-5

SYSTEMATIC REVIEW

Assessing Patient Risk, Benefit, and Outcomes in Drug Development: A Decade of Lenvatinib Clinical Trials: A Systematic Review

Patrick Crotty[1] · Karim Kari[1] · Griffin K. Hughes[1] · Chase Ladd[1] · Ryan McIntire[1] · Brooke Gardner[1] · Andriana M. Peña[1] · Sydney Ferrell[5] · Jordan Tuia[4] · Jacob Cohn[3] · Alyson Haslam[4] · Vinay Prasad[4] · Matt Vassar[1,2]

Accepted: 2 February 2024 / Published online: 11 March 2024
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2024

Abstract

Importance Chemotherapy agents are typically initially tested in their most promising indications; however, following initial US FDA approval, new clinical trials are often initiated in less promising indications where patients experience a worse burden-benefit ratio. The current literature on the burden-benefit profile of lenvatinib in non-FDA-approved indications is lacking.

Objective This study aimed to evaluate published clinical trials of lenvatinib in order to determine the burden-benefit profile for patients over time.

Evidence Review On 25 May 2023, we searched the Pubmed/MEDLINE, Embase, Cochrane CENTRAL, and ClinicalTrials.gov databases for clinical trials of lenvatinib used to treat solid cancers. Eligible articles were clinical trials, containing adult participants, published in English, and involving solid tumors. Screening and data collection took place in a masked, duplicate fashion. For each eligible study, we collected adverse event data, trial characteristics, progression-free survival (PFS), overall survival (OS), and objective response rate (ORR). Trials were classified as positive when meeting their primary endpoint and safety, negative (not meeting either criteria), or indeterminate (lacking prespecified primary endpoint).

Findings Expansion of clinical trial testing beyond lenvatinib's initial FDA indication demonstrated a consistent rise in cumulative adverse events, along with a decline in drug efficacy. Lenvatinib was tested in 16 cancer indications, receiving FDA approval in 4. A total of 5390 Grade 3–5 adverse events were experienced across 6225 clinical trial participants. Expanded indication testing further demonstrated widely variable ORR (11–69%), OS (6.2–32 months), and PFS (3.6–15.7 months) across all indications. After initial FDA approval, clinical trial results in expanded indications were less likely to meet their primary endpoints, particularly among non-randomized clinical trials.

Conclusion and relevance Our paper evaluated the effectiveness of lenvatinib for its FDA-approved indications; however, expansion of clinical trials into novel indications was characterized by diminished efficacy, while patients experienced a high burden of adverse events consistent with lenvatinib's established safety profile. Furthermore, clinical trials testing in novel indications was marked by repeated phase I and II clinical trials along with a failure to progress to phase III clinical trials. Future clinical trials using lenvatinib as an intervention should carefully evaluate the potential benefits and burden patients may experience.

1 Introduction

Cancer drug development and marketing has been characterized by increasing costs over the past decade. Difficulty in recruiting patients, demand for technical expertise, stringent regulatory requirements, and lengthy trial durations are often cited as key drivers of high drug development costs [1, 2]. However, decrements in quality of life from adverse effects are often not considered in the economic evaluation of cancer drug therapies [3]. In addition, the average time for a new trial drug to progress from the start of clinical testing to US FDA approval is 7.2 years [1]. Throughout this process, pharmaceutical companies incur a substantial cost averaging $1.6–$2.5 billion, including cost associated with failed drug development [1, 2, 4]. It is estimated that 90% of novel drugs fail to reach FDA approval, thus the financial burden associated with failed trials is concerning [5]. Therefore, importance should be placed on optimizing outcomes and enhancing research to minimize risk associated with costly and unnecessary drug trials.

Extended author information available on the last page of the article

Key Points

Lenvatinib clinical trials testing novel expanded indications demonstrated variable objective response rates, overall survival, and progression-free survival, while being less likely to reach their primary endpoints compared with those indications receiving US FDA approval.

Lenvatinib clinical trials in expanded indications failed to progress to later-stage clinical trials and subsequent regulatory approval.

Post-FDA approval, lenvatinib trials for expanded indications showed a rising adverse event rate and declining efficacy, emphasizing the need to reconsider conducting trials beyond the original drug design.

To improve the cost, impact, and burden of cancer treatments, researchers should strive to improve therapeutic benefits of treatments [6]. Two separate studies conducted by Carlisle et al. examined sunitinib and imatinib, revealing a worsening risk/benefit ratio when explored for expanded indication use [7, 8]. For example, imatinib initially demonstrated positive results, with 60% of trials launched later receiving FDA approval within 8 years [8]. However, after initial approval, only 6% of subsequent trials managed to either gain FDA approval or show positive results within the same time frame [8]. Similarly, early trials of sunitinib showed favorable outcomes in treating both gastrointestinal stromal tumors and renal cell carcinoma (RCC); however, the burden/benefit profile deteriorated when tested in new indications [7]. These findings suggest therapies driven by molecular insight support their primary indication, while expanded indication trials in novel malignancies face a diminished benefit-risk profile. [9, 10]

Lenvatinib is a tyrosine kinase inhibitor manufactured as Lenvima®, which inhibits vascular endothelial growth factor (VEGF) receptors VEGFR1 (FLT1), VEGFR2 (KDR), and VEGFR3 (FLT4) activity [11]. In 2015, the FDA approved lenvatinib for patients with radioactive iodine-refractory differentiated thyroid cancer (DTC) [12]. Since then, lenvatinib has received additional FDA approval in hepatocellular carcinoma (HCC), RCC, and endometrial carcinoma (EC) [11]. However, as lenvatinib trials have expanded across multiple indications, the cumulative patient benefit and safety risks have not been characterized. Redundant studies and continued exploration for new drug indications increase financial burden on patients and healthcare [13]. Furthermore, as additional clinical trials for new indications are conducted, the risk-benefit balance for patients may shift over time. Lenvatinib was selected for this study due to its recent FDA approvals, the variety of solid tumor indications under exploration, and the need to better understand cumulative adverse events and efficacy over time. Therefore, the goal of this study is to assess the total patient benefit and burden for lenvatinib by creating a catalog of published clinical trials. Furthermore, we anticipate the results of this study will lend valuable insight into the implications of cancer treatment for healthcare providers, patients, and policymakers.

2 Methods

2.1 Study Design/Open Science

This was a cross-sectional study that investigated the risk/benefit profiles of lenvatinib (*Lenvima*, Eisai Inc.) clinical trials during their development and applications to indications outside of their initial approval. For the enhancement of rigor, reproducibility, and open science, we uploaded a protocol *a priori*. Upon the completion of the investigation, we uploaded raw data, statistical analysis scripts, and extraction forms to Open Science Framework (OSF), a free-to-upload data repository [14]. The data will remain available on OSF throughout the repository's lifecycle or upon request (https://osf.io/vdr68/?view_only=78fb8b6f2e3e440aa71423b40ec8dd46) [15].

2.2 Research Questions, Definitions, and Hypothesis

Given that clinical trials are costly and potentially harmful to patients, what are the benefit/risk profiles of clinical trials assessing the efficacy of lenvatinib? Do the combined risk profiles—the drug's burden/benefit portfolio—represent an overall excessive risk to patients? A clinical trial *profile* was defined as the total risk and benefit that participants experienced during a single trial, measured through selected tools as mentioned in the *Data Extraction* section. A drug *portfolio* was defined as the complete collection of trial profiles for a specific intervention. It was hypothesized that the expansion of clinical trials of lenvatinib into expanded indications would result in an overall negative drug portfolio with more negative trials and an increased risk to patients.

2.3 Literature Search

A literature search of the Pubmed/MEDLINE, Embase (Elsevier), Cochrane CENTRAL, and ClinicalTrials.gov databases was performed on 25 May 2023 to explore cancer treatment clinical trials that used lenvatinib as monotherapy or in combination with other interventions. Our search strings were standardized across all databases using

the PolyGlot Search Translator (https://sr-accelerator.com/#/polyglot), which was developed by Bond University and the Institute for Evidence-Based Healthcare [16]. The search strings, including the date of search and initial returns, were uploaded to OSF and were available as supplementary data in the final manuscript submission.

2.4 Selection Process

Search returns were uploaded to Rayyan for literature screening. Two authors (PC and KK) screened the titles and abstracts in a masked duplicate procedure for potential inclusion. Any discrepancies were resolved by a third author (CL) after the screening was completed. Reasons for exclusion were documented during the screening process to form a study exclusion flowchart.

2.5 Inclusion and Exclusion Criteria

Inclusion criteria required the study to (1) be a clinical trial of adult human subjects; (2) evaluate the efficacy of lenvatinib as an intervention for treating solid cancerous tumors in adults; (3) assess the benefit of lenvatinib using objective response rate (ORR) according to the Response Evaluation

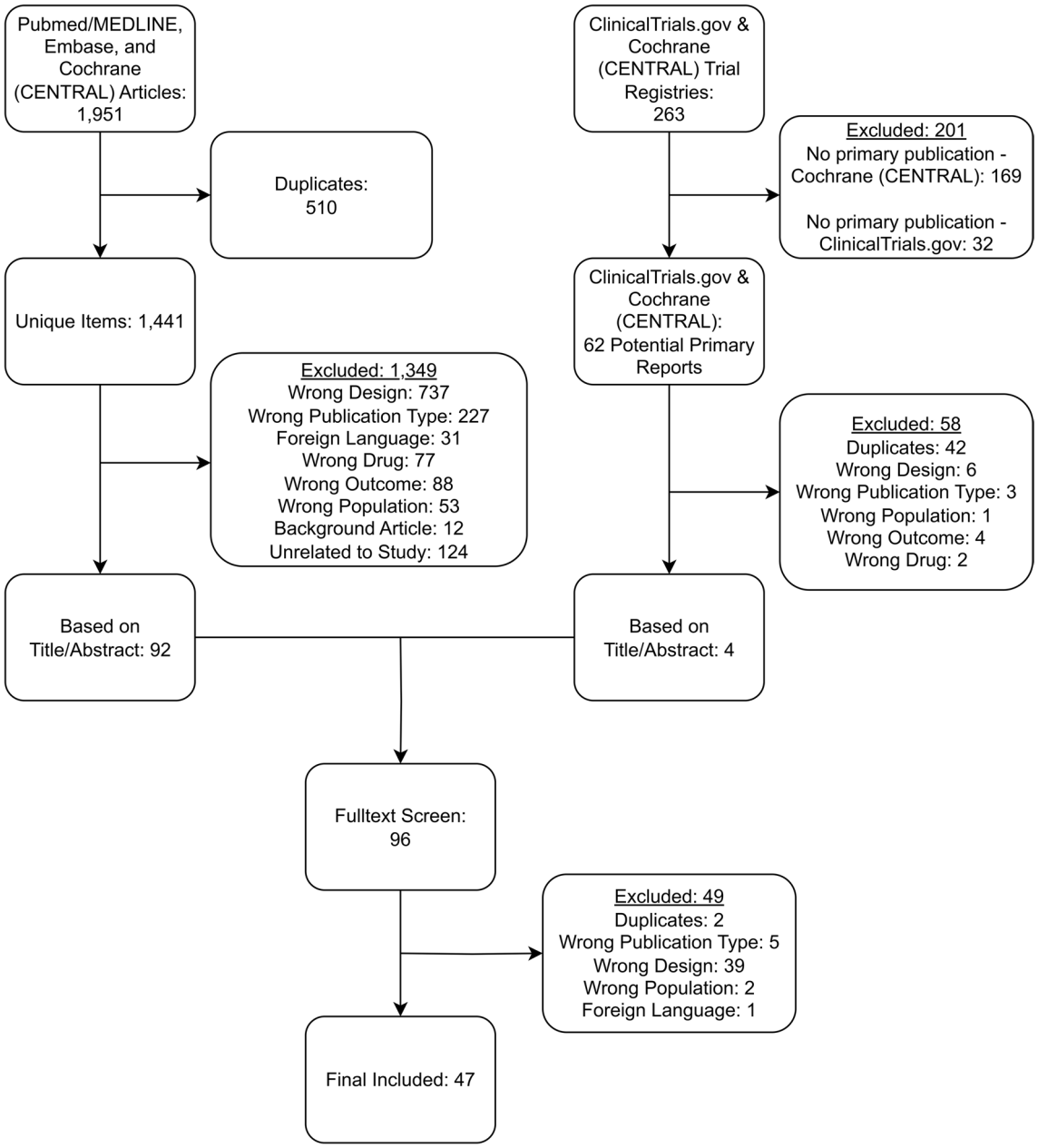


Fig. 1 Flow diagram for study inclusion. Forty-seven studies were ultimately included for analysis.

Table 1 Characteristics of the included trials

Characteristic	Overall [$n = 47$]	Combination [$n = 25$]	Monotherapy [$n = 22$]
Phase			
1	11 (23.4)	8 (32.0)	3 (13.6)
2	29 (61.7)	13 (52.0)	16 (72.7)
3	7 (14.9)	4 (16.0)	3 (13.6)
Stage			
Metastatic	36 (76.6)	21 (84.0)	15 (68.2)
Non-metastatic	11 (23.4)	4 (16.0)	7 (31.8)
Response criteria used			
RECIST	40 (85.1)	20 (80.0)	20 (90.9)
mRECIST	5 (10.6)	3 (12.0)	2 (9.1)
Other	2 (4.3)	2 (8.0)	0 (0.0)
Results			
Positive	29 (61.7)	13 (52.0)	16 (72.7)
Indeterminate	10 (21.3)	7 (28.0)	3 (13.6)
Negative	8 (17.0)	5 (20.0)	3 (13.6)
Randomized			
Non-randomized	36 (76.6)	18 (72.0)	18 (81.8)
Randomized	11 (23.4)	7 (28.0)	4 (18.2)
Randomization ratio			
1:1	6 (60.0)	4 (66.7)	2 (50.0)
1:1:1	2 (20.0)	2 (33.3)	0 (0.0)
2:1	2 (20.0)	0 (0.0)	2 (50.0)
Unknown/non-randomized	37	19	18
Blinding			
Non-blinded	44 (93.6)	25 (100.0)	19 (86.4)
Double	3 (6.4)	0 (0.0)	3 (13.6)
Number of centers			
Multicenter	29 (61.7)	13 (52.0)	16 (72.7)
Single center	14 (29.8)	11 (44.0)	3 (13.6)
Not stated	4 (8.5)	1 (4.0)	3 (13.6)
Sponsor/funding			
Industry	31 (66.0)	15 (60.0)	16 (72.7)
Industry, non-industry	4 (8.5)	3 (12.0)	1 (4.6)
Government	3 (6.4)	3 (12.0)	0 (0.0)
Government, non-industry	3 (6.4)	1 (4.0)	2 (9.1)
Industry, Government	3 (6.4)	2 (8.0)	1 (4.6)
Industry, Government, non-industry	3 (6.4)	1 (4.0)	2 (9.1)
Conflicts of interest/disclosure statement			
Reports conflicts of interest	37 (78.7)	18 (72.0)	19 (86.4)
Reports no conflicts of interest	8 (17.0)	6 (24.0)	2 (9.1)
Not reported	2 (4.3)	1 (4.0)	1 (4.6)
Country			
United States	15 (31.9)	11 (44.0)	4 (18.2)
Japan	13 (27.7)	4 (16.0)	9 (40.9)
China	9 (19.2)	6 (24.0)	3 (13.6)
France	2 (4.3)	0 (0.0)	2 (9.1)
Spain	2 (4.3)	1 (4.0)	1 (4.6)
Taiwan	2 (4.3)	2 (8.0)	0 (0.0)
Belgium	1 (2.1)	0 (0.0)	1 (4.6)
Italy	1 (2.1)	0 (0.0)	1 (4.6)
Netherlands	1 (2.1)	0 (0.0)	1 (4.6)
Singapore	1 (2.1)	1 (4.0)	0 (0.0)

Data are expressed as *n* (%)

RECIST Response Evaluation Criteria in Solid Tumors, *mRECIST* modified RECIST

Criteria in Solid Tumors (RECIST) criteria; and (4) be published in English. Non-oncological studies, studies on non-solid tumors, studies on biosimilars, pharmacology studies on healthy participants, and studies involving exclusively pediatric populations were excluded. Also excluded were secondary reports, interim results, clinical trial updates and follow-ups, preclinical studies, literature reviews, systematic reviews, meta-analyses, human tissue studies, laboratory studies, case reports, letters to the editor, editorials, opinion pieces, conference abstracts, corrections, or redactions. Studies with manuscripts written in a language other than English were also excluded.

2.6 Data Extraction

Following data screening, the final study pool was subjected to a data extraction process, which was carried out by two authors (PC and KK) in a masked, duplicate manner, while a third author (CL) resolved discrepancies. The authors recorded the following variables: published trial title, PubMed ID, clinical trial registry number, country of first author's affiliation, date of publication, number of participants, mean or median age of participants, number of male participants, number of female participants, indications of the trial, metastatic or non-metastatic stage, whether the trial was controlled, if the trial assessed monotherapy or combination therapies, phase of the trial, number of centers, blinding of trial participants, randomization ratio, and study sponsor, including funding and conflicts of interest statements.

As clinical trials progress to later stages, primary and secondary endpoints evolve from proving basic safety to efficacy as regulatory approval and endorsement is sought. The following variables were extracted for treatment arms to measure risks and benefits: arm name, Common Terminology Criteria for Adverse Events (CTCAE) grade of adverse events, median progression-free survival (PFS) in months, hazard ratio (HR) of PFS, median overall survival (OS) in months, HR of OS, partial response (PR), complete response (CR), ORR as defined in the RECIST criteria, number of grade 3–5 adverse events, dose reductions, discontinuations due to drug-related adverse events, and whether the trial was positive, indeterminate, or negative. Data on outcome measurements and adverse events sustained by all trial participants of a prespecified indication were recorded. A trial was considered positive if it utilized a tolerable regimen and met its prespecified endpoints. An indeterminate trial did not prespecify endpoints and was using a tolerable regimen. A trial was considered negative if it failed to meet its prespecified endpoints or did not utilize a tolerable regimen. All phase I trials were classified as indeterminate rather than positive or negative to more accurately reflect their unique characteristics [17]. The tolerability of a regimen was determined by the authors of the trial [8].

Several design decisions were implemented for trial-specific characteristics. In instances where trials reported multiple phases, the higher phase was extracted. If a trial provided a response rate without specifying the proportions of partial or CRs, it was assumed that only PR was measured. Extracted data focused on confirmed responses when trialists specified them as confirmed or unconfirmed. Additionally, independently confirmed measurements were extracted if trialists indicated that measurement confirmation was conducted by independent investigators. Trials incorporating two or more different indication types were categorized as 'multiple indications', with additional clarification available. ORRs were calculated for all participants in a specific arm, unless the trialist specified evaluable patients.

2.7 Statistical Analysis

Descriptive statistics were performed using R version 4.2.1 and RStudio (The R Foundation for Statistical Computing, Vienna, Austria). The overall response rate (ORR) was calculated for each treatment arm by dividing the sum of CRs and PRs by the total number of participants. ORRs for each indication were then pooled by finding the median ORR across all ORR arms within that indication. The cumulative adverse event rate was determined by dividing the total number of grade 3–5 adverse events by the total number of enrolled participants. Based on reviewer feedback, trend lines by phase were visualized by monotherapy or combination therapy using the techniques described by Mattina et al. [18] The protocol was amended to reflect these changes, which is viewable on the OSF.

A random effect meta-analysis of raw proportions, using a restricted maximum likelihood estimator to pool ORRs across trial characteristics, was conducted using R version 4.2.1 and RStudio. After calculations, we axis transferred the proportions to improve readability.

2.8 Ethical Oversight

The Oklahoma State University Center for Health Sciences reviewed our protocol and determined that this research qualifies as non-human subject research as defined in regulation 45 CFR 46.102(d) and (f).

3 Results

3.1 General Characteristics

Our systematic search yielded 1951 studies cataloged in bibliographic databases. Our clinical trial registry search

Table 2 Trial characteristics and outcomes by indication for (a) monotherapy, (b) combination therapy

(a)

Solid tumor indications	Trajectory launched year	No. of trials	No. of randomized trials [*n* (%)]	No. of participants	No. of Grade 3–5 events	Dose reductions	Dose D/Cs	Median PFS (months)	Median OS (months)	Median partial response rate (%)	Median complete response rate (%)	Median ORR[a] (%)
Multiple indications	2005	2	0 (0)	159	56	23[c]	17	17	–	10.1	0.0	10.1
Differentiated thyroid cancer	2008	4	3 (75)	753	514	222	52	15.45	–[b]	50.0	0.0	50.0
Thyroid cancer	2008	4	0 (0)	169	139	130	23	9	12.1	23.5	0.0	23.5
Hepatocellular carcinoma	2009	3	1 (25)	1025	764	217	53	5.6	13.6	16.7	0.0	16.7
Endometrial cancer	2010	1	0 (0)	133	111	38	24	5.6	10.6	13.5	0.8	14.3
Non-small cell lung cancer	2013	1	0 (0)	25	58	16	6	6.4	–	16.0	0.0	16.0
Adenoid cystic carcinoma	2015	2	0 (0)	60	53	47	19	13.3	27	13.5	0.0	13.6
Biliary tract cancer	2015	2	0 (0)	67	37	20	2	2.7	9.4	11.8	0.0	11.9
Gastroenteropancreatic neuroendocrine tumors	2015	1	0 (0)	111	76	90	16	15.7	32	28.9	0.0	28.9
Colorectal cancer	2016	1	0 (0)	30	25	27	2	3.6	7.4	6.7	0.0	6.7
Thymic carcinoma	2017	1	0 (0)	42	52	42	7	9.3	–	38.1	0.0	38.1
Totals/medians		22	4 (18.2)	2574	1885	872	221	9.1	12.2	24.5	1.1	25.6

(b)

Solid tumor indications	Trajectory launched year	No. of trials	No. of randomized trials [*n* (%)]	No. of participants	No. of Grade 3–5 events	Dose reductions	Dose D/C's	Median PFS (months)	Median OS (months)	Median partial response rate (%)	Median complete response rate (%)	Median ORR[a] (%)
Melanoma	2005	2	0 (0)	135	61	60	11	4.2	14	18.6	0.0	20.1
Non-small cell lung cancer	2009	1	0 (0)	28	88	–[c]	1	9	–	57.1	0.0	60.7
Renal cell carcinoma	2010	7	4 (57.1)	1768	1845	1140	338	11.1	22	34.0	0.0	36.1
Hepatocellular carcinoma	2013	5	2 (28.6)	609	471	267	52	8.9	17.1	33.4	0.0	35.0

Table 2 (continued)

(b)

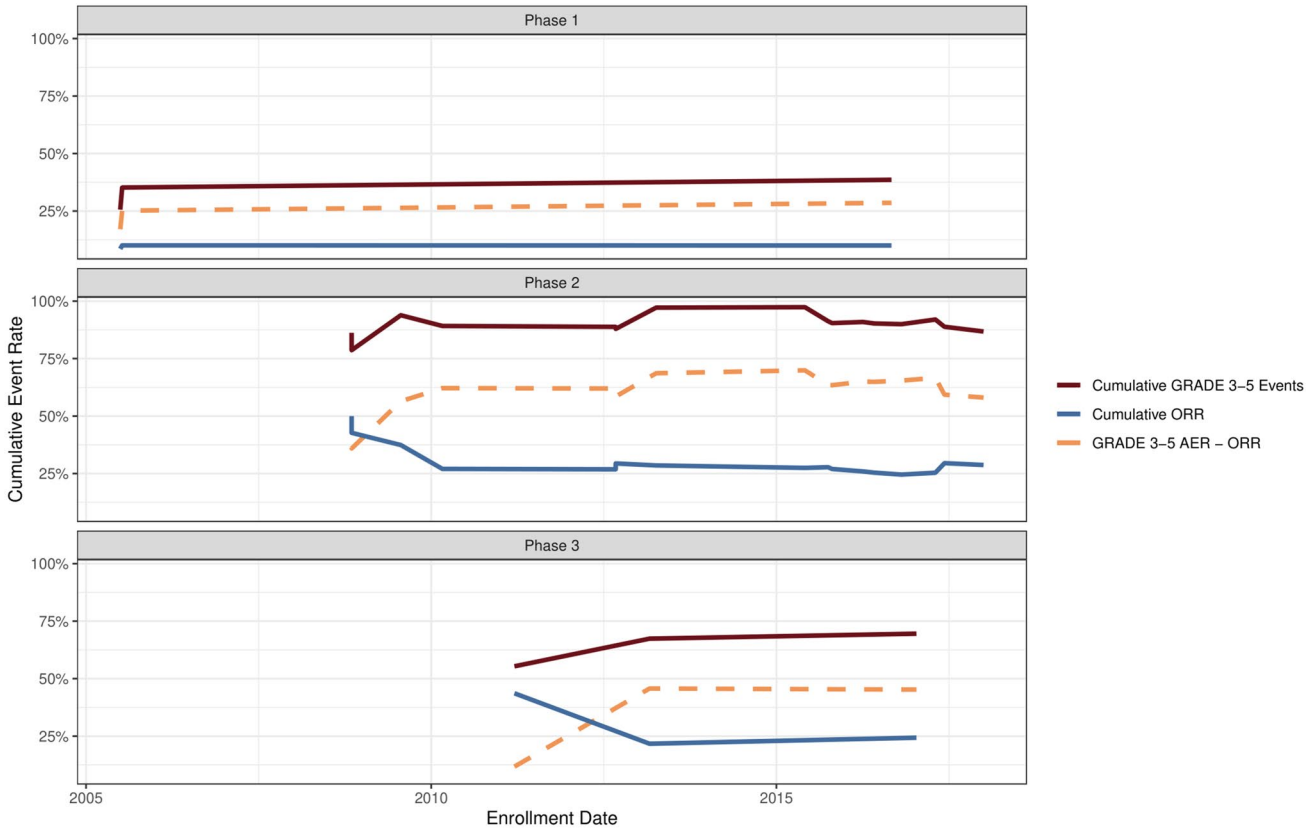
Solid tumor indications	Trajectory launched year	No. of trials	No. of randomized trials [*n* (%)]	No. of participants	No. of Grade 3–5 events	Dose reductions	Dose D/C's	Median PFS (months)	Median OS (months)	Median partial response rate (%)	Median complete response rate (%)	Median ORR[a] (%)
Multiple indications	2015	3	0 (0)	169	188	135	51	5.7	–	41.4	8.3	49.3
Breast cancer	2015	1	0 (0)	47	43	6	–	–[b]	–	34.8	0.0	34.8
Biliary tract cancer	2018	3	0 (0)	90	59	11[c]	3	10	17.7	45.5	0.0	45.5
Endometrial cancer	2018	1	1 (14.3)	827	714	320	165	5.5	14.9	18.7	0.0	23.3
Gastric cancer	2018	1	0 (0)	29	32	29	0	7.1	–	65.5	0.0	69.0
Squamous cell carcinoma	2021	1	0 (0)	14	4	–	–	4.6	6.2	28.6	0.0	28.6
Total/medians		25	7 (100)	3716	3505	1970	621	9.2	17.7	33.5	16.3	49.8

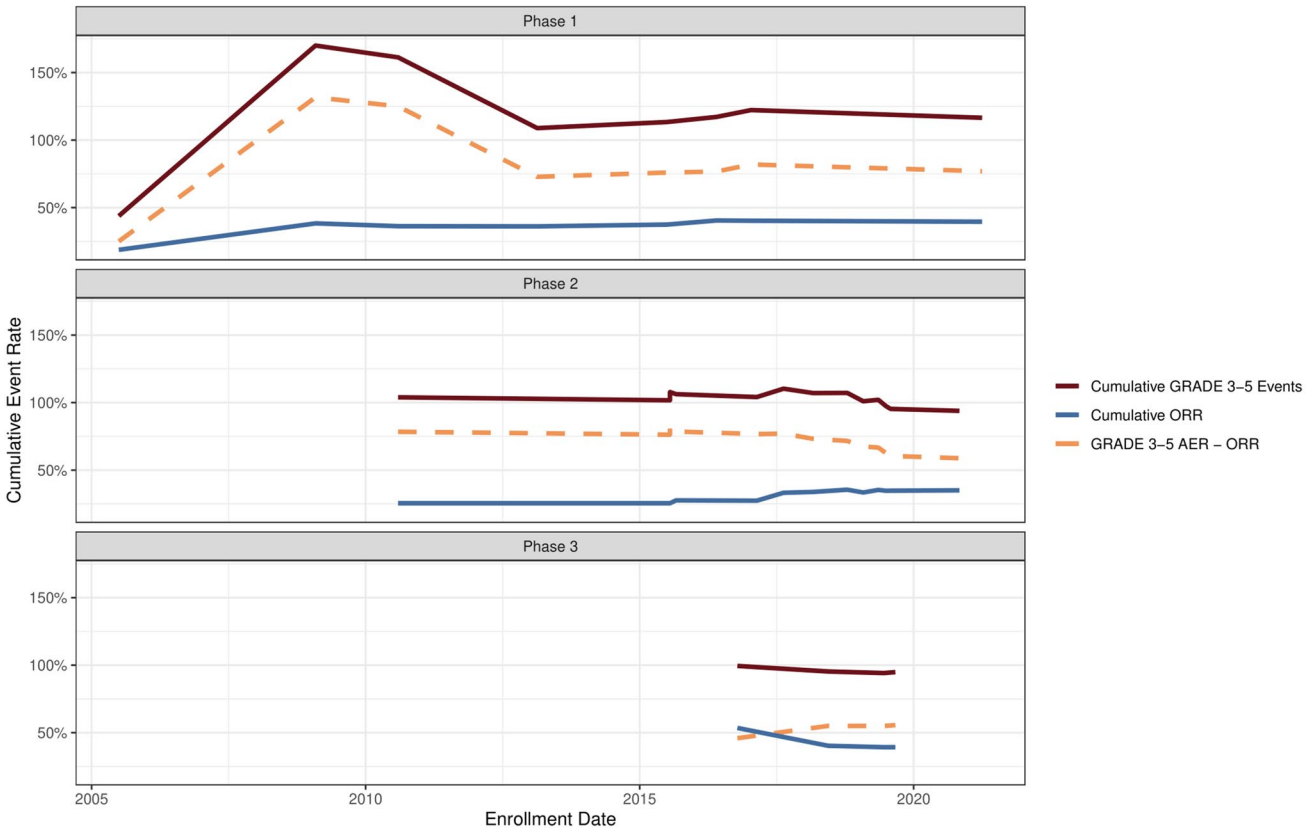
PFS progression-free survival, *OS* overall survival, *ORR* objective response rate, *RECIST* Response Evaluation Criteria in Solid Tumors, *mRECIST* modified RECIST, *D/Cs* dose discontinuations

[a]Response rate median calculations for those indications using RECIST or mRECIST only

[b]Did not measure outcome, failed to report outcomes, or failed to report outcome measure for all enrolled patients

[c]Not all data were reported


Phase 1
Phase 2
Phase 3
100%
75%
50%
25%
Cumulative Event Rate
2005
2010
2015
Enrollment Date
Cumulative GRADE 3–5 Events
Cumulative ORR
GRADE 3–5 AER – ORR


Phase 1
Phase 2
Phase 3
150%
100%
50%
Cumulative Event Rate
2005
2010
2015
2020
Enrollment Date
Cumulative GRADE 3–5 Events
Cumulative ORR
GRADE 3–5 AER – ORR

◂**Fig. 2** (**a**) Cumulative AERs of monotherapy trials by enrollment date versus cumulative ORR separated by phase. (**b**) Cumulative AERs of combination therapy trials by enrollment date versus cumulative ORR separated by phase. *Δ[AER-ORR]* represents the absolute difference between cumulative AER and cumulative ORR, *ORR* objective response rate

yielded an additional 263 studies. Following a full-text screen of 96 publications, 49 studies were excluded, resulting in 47 studies for inclusion (Fig. 1). Of our sample of 47 studies, 22 (46.8%) used lenvatinib as a monotherapy and 25 (53.2%) used combination therapy. These clinical trials involved a total of 6290 participants. Lenvatinib has received FDA approval in four indications; clinical trials were conducted across 16 indications in our sample. The most common indications for lenvatinib were HCC (8/47, 17.0%) and RCC (7/47, 14.9%). Forty-four (93.6%) studies were unblinded, while 3 (6.4%) studies were double-blinded. Thirty-six studies were not randomized (76.6%), while 11 (23.4%) were randomized. Positive results were found in 29 (61.7%) studies, 10 (21.3%) were indeterminate, and 8 (17%) yielded negative results. Additional trial characteristics can be found in Table 1.

3.2 Endpoints

Among FDA-approved indications, HCC, lenvatinib monotherapy trials demonstrated a median PFS of 5.6 months, a median OS of 13.6 months, and a median ORR of 16.7%. In contrast, lenvatinib HCC combination therapy trials resulted in a median PFS of 8.9 months, a median OS of 17.1 months, and an ORR of 35%. In the treatment of endometrial cancer, monotherapy demonstrated a median PFS of 5.6 months, a median OS of 10.6 months, and an ORR of 14.3%. Meanwhile lenvatinib combination therapy yielded a median PFS of 5.5 months, a median OS of 14.9 months, and an ORR of 23.3%. These findings underscore the promising impact of combination strategies on patient outcomes within certain indications. When comparing the overall OS and ORR between monotherapy and combination therapy, it is evident that combination therapy demonstrated an elevated OS (17.7 months) and ORR (49.8%) in contrast to monotherapy (12.2 months, 25.6%). Meanwhile, median PFS was similar between the overall combination and monotherapy groups (Table 2b).

3.3 Risk Assessment

In our cumulative analysis of lenvatinib trials, a total of 6290 participants and 5390 grade 3–5 adverse effects were

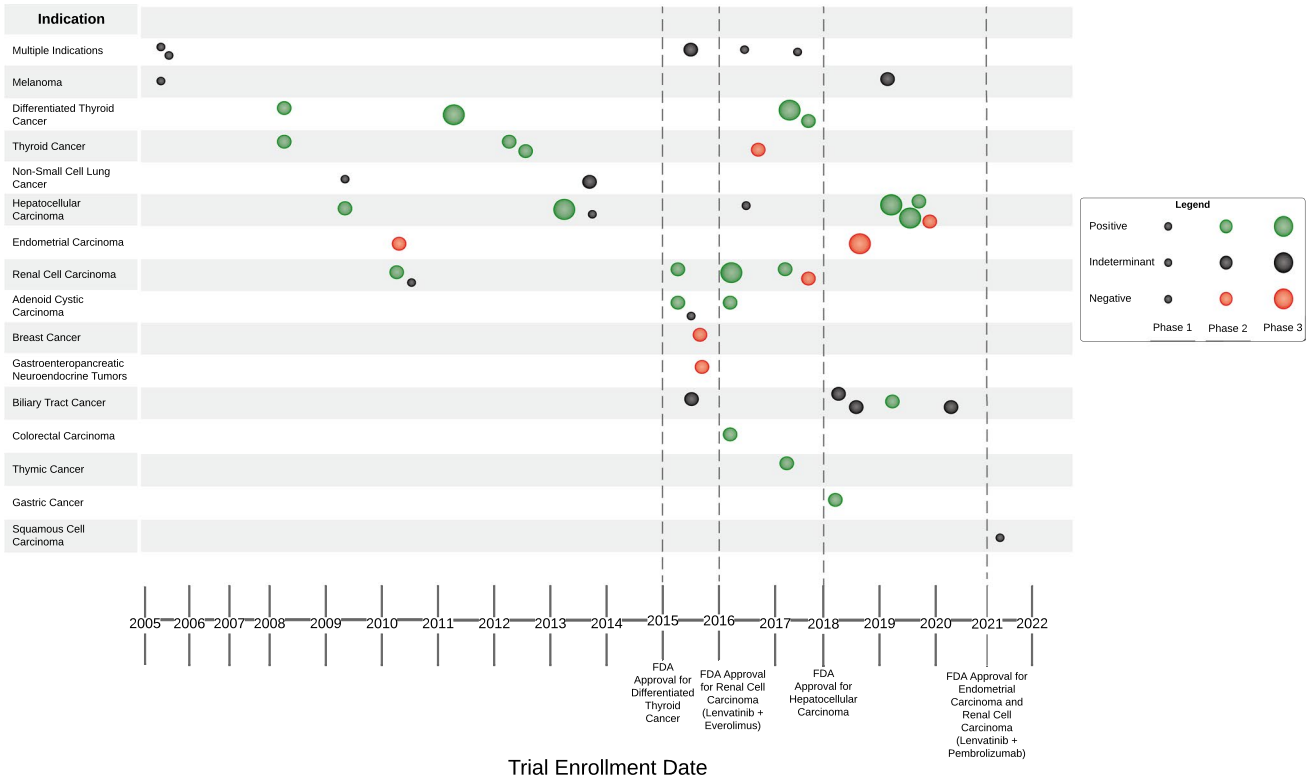


Fig. 3 Lenvatinib trials visualized by indication and date of publication, with each point representing a trial's phase and relative participant number. Blue, red, or gray points represent positive, negative, or indeterminate results, and varying sizes denote the phase

experienced across these 47 clinical trials. The cumulative nature of these estimates reflects the evolving landscape of clinical trials over time, accounting for both an increase in the number of

trials and their growing sizes. Furthermore, the total participants and grade 3–5 adverse events include all comparator arms, thus capturing the nature of lenvatinib clinical trials as a whole. Only 12 trials (25.5%) reported all adverse events. Notably, the year 2013 recorded an adverse event rate (AER) >100%, indicating more adverse events were recorded than participating patients. Additionally, although all indications were off-label at one time before getting approved, the expansion of lenvatinib clinical trials into expanded indications was also associated with another spike in adverse events (electronic supplementary material [ESM] A).

Each specific phase of lenvatinib clinical trials maintained a relatively consistent rate of ORR and AER. Our study used Grade–ORR, the difference between Grade 3–5 AER and ORR, to measure the burden-benefit profile of lenvatinib, with an increasing Grade–ORR indicating a worsening burden-benefit profile. Phase II and III clinical trials generally demonstrated a Grade–ORR just above 50%. The phase I monotherapy outliers exhibited a Grade–ORR approximately half that of phase II and III trials, while phase 1 combination outliers reported a slightly higher Grade–ORR (Fig. 2a). Additionally, the smaller numbers of participants and limited number of clinical trials extracted make these results more susceptible to outliers. Among clinical trial participants, 74% of monotherapy and 86% of combination therapy clinical trial participants were in an indication that received FDA approval (Table 2). This predominance stabilized the cumulative rates of ORR and Grade 3–5 over time.

The rates of adverse events were consistently in excess of 75% of participants among the majority of indications; however, the ORR among these participants was variable and the number of indications that followed a standard developmental path from phase I through III were limited to DTC, RCC, HCC, and EC.

3.4 Accumulating Evidence and Research Organization (AERO) Diagram

Figure 3 illustrates the growth and progression of the lenvatinib clinical trial portfolio, visualizing various phases and indications [19]. The stark contrast in trial outcomes becomes evident when comparing trials conducted for FDA-approved indications with those conducted for expanded indications. Among the 20 trials conducted for FDA-approved indications, 65% were positive (13/20), whereas for trials conducted for expanded indications, only 33% (9/27) of the trials produced positive results. Thus, it becomes evident that when lenvatinib is utilized for FDA-approved indications, it consistently exhibits efficacious outcomes; however, when used for expanded indications, the observed benefit decreases, while the frequency of adverse events remains high. A significant example of a negative result is the pivotal phase III clinical trial KEYNOTE-775, conducted in EC. This trial played a crucial role in securing FDA approval for lenvatinib in this indication. Although demonstrating efficacy with an ORR of 31.9%, the trial failed to meet its primary endpoint of PFS, as indicated by an HR of 0.56 (trialist prespecified HR of 0.55), ultimately leading to a negative outcome. [20]

3.5 Meta-Analysis

ORRs were pooled through raw proportion, random effect meta-analysis for monotherapy phase I (9.81%, 95% confidence interval [CI] 5.47–14.15%, $I^2 = 0.0\%$, $\tau^2 = 0.00$), phase II (24.41%, 95% CI 17.11–31.71%, $I^2 = 86.4\%$, $\tau^2 = 0.02$), and phase III (34.44%, 95% CI 12.62–56.26%, $I^2 = 98.5\%$, $\tau^2 = 0.04$) clinical trials. The same analyses were conducted for combination therapy phase I (41.81%, 95% CI 28.24–55.38%, $I^2 = 75.9\%$, $\tau^2 = 0.03$), phase II (39.84%, 95% CI 29.02–50.67%, $I^2 = 90.9\%$, $\tau^2 = 0.03$), and phase III (37.25%, 95% CI 24.23–50.26%, $I^2 = 97.4\%$, $\tau^2 = 0.02$) trials. Further meta-analyses were conducted of ORR by indication (30.79%, 95% CI 23.50–38.07%, $I^2 = 97.0\%$, $\tau^2 = 0.02$) and among all clinical trials (31.67%, 95% CI 26.47–36.88%, $I^2 = 95.0\%$, $\tau^2 = 0.03$). Forest plots of each meta-analysis may be found in the ESM.

4 Discussion

4.1 General Findings

Pharmaceutical companies frequently pursue additional FDA approvals for drugs, extending beyond the initially approved indications [7]; however, it is noteworthy that clinical trials in novel indications are less likely to secure FDA approval compared with trials conducted before the drug's initial FDA approval [10]. A concerning trend our study noted was the failure of clinical trial testing in novel indications to advance to phase III clinical trials. In the case of FDA-approved indications, all trials eventually progressed to a phase III trial, indicating a standardized progression of clinical development. However, it is worth noting that for the expanded indications, the progression from phase I to phase III trials was absent. Instead, these indications often underwent repeated phase I and II trials without further advancement to phase III trials. For example, biliary tract cancer underwent five phase II trials without ever leading to phase III trials. Additionally, clinical trials in non-FDA indications were less likely to have achieved a positive outcome compared with those

in indications that would receive FDA approval (33% vs. 65%). This pattern was unexpected because the exploration of novel indications in clinical trials primarily involved early phase, often non-randomized clinical trials, which are more prone to overstating efficacy benefits such as ORR. These discrepancies highlight the variation in the development and evaluation trajectories for different indications of lenvatinib.

Sunitinib and imatinib are cancer treatment drugs characterized by initially successful treatment in approved indications with worsening expanded indication use [7, 8]. A 2021 study by Motzer et al. demonstrated superior results for lenvatinib compared with sunitinib in a randomized phase III trial in advanced RCC [21]. Lenvatinib combination arms demonstrated higher ORR and PFS but an OS rate was not reached by any group [21]. Lenvatinib is characterized by a safety profile where the number of Grade 3–5 adverse events to participant ratio is often in excess of 75% of participating patients and is relatively consistent across indications. Worsening expanded indication efficacy with consistently high rates of adverse events suggests lenvatinib is potentially following the trajectory of sunitinib and imatinib. A worsening benefit/burden ratio needlessly places clinical trial participants at risk for severe adverse events with diminishing therapeutic benefit.

Of the 16 potential indications, lenvatinib has received FDA approval for four of its indications, with ORR values ranging from 18.8 to 71%. In the context of endometrial cancer, two studies have been conducted. In 2020, Vergote et al. conducted a single-arm, phase II study examining lenvatinib as a monotherapy in advanced endometrial cancer [22]. After trial completion, they found an ORR value of 14.3%. A follow-up study conducted in 2022 by Makker et al. compared lenvatinib and pembrolizumab with physician choice chemotherapy [20]. Their results demonstrated an ORR rate of 31.9% with combination therapy. Subsequently, the study by Makker et al. led to the FDA approval of lenvatinib with pembrolizumab for endometrial cancer. Despite its approval for this indication, our study found that the median ORR (14.7%) was vastly lower relative to the other FDA-approved indications. This is due to a small number of endometrial studies and differences in values among the combination and monotherapy study arms.

4.2 Adverse Events

A potent tyrosine kinase inhibitor, lenvatinib's safety profile consistently exhibits high rates of adverse events among clinical trial participants. The most efficacious phase III randomized studies of lenvatinib, used to grant FDA approval in several indications, exhibited AER in excess of 75%, with comparatively high ORR (Fig. 2a, b). In the years following initial FDA approval, as the cumulative adverse event rate of lenvatinib rose over time, the cumulative ORR declined. Potent chemotherapies often justify their potential adverse effects because of the therapeutic benefit they can offer patients [23]. However, with the expansion of lenvatinib clinical trials into novel indications, we noticed a continued increase in patient burden with a decreasing cumulative ORR over time. The initial decrease in cumulative ORR was surprising considering the abundance of non-randomized trials and their tendency to often overstate efficacy values such as ORR by up to 2.5 times compared with randomized controlled trials [24]. In 2021, the clinical trial by Motzer et al. temporarily reversed the downward trend, as the large sample size led to a substantial ORR increase and corresponding AER spike [21]. In the years following this study, we noticed another initial decrease and then plateauing of cumulative ORR. Hutchinson et al. found a declining probability in FDA approval in novel cancer indications in the years after initial FDA approval [25]. A diminished likelihood for regulatory approval, coupled with high rates of adverse events and declining drug efficacy suggest a necessary review of the burden and benefits patients may experience in novel clinical trial testing. [25]

4.3 Strengths and Limitations

Our study has several strengths. First, to facilitate reproducibility and transparency, we uploaded our protocol, data extraction forms, and data sheets to OSF, a free open-source archive. In addition, investigators completed training sessions to standardize data extraction. Lastly, our data were extracted in a masked, duplicated fashion, to further enhance the validity of our study. This practice is considered the gold standard for data extraction as set forth by Cochrane Collaboration guidelines [26]. However, our study, although methodologically rigorous, comes with limitations. Errors in data extraction could have persisted into our final sample analysis, however we employed Cochrane Collaboration guidelines to mitigate this risk. Our study is also cross-sectional by design, and thus these findings are not generalizable to other cancer drugs or fields of medicine. A further limitation to our study is only analyzing peer-reviewed published clinical trials, potentially excluding some clinical trials from our sample. Additionally, further limitations include limited follow-up time for more recent trials, challenges pooling outcomes across disparate studies, and difficulties quantifying the overall value/impact of individual trajectories.

5 Conclusions

Our paper analyzed the burden and benefits of clinical trials that patients experience, when investigators attempt to expand the clinical use of lenvatinib beyond its

FDA-approved indications. The adverse event rates across indications remained considerable, while lenvatinib's benefits were diminished in non-FDA-approved indications. Furthermore, clinical trial testing in expanded indications was characterized by absent progression to phase III clinical trials and instead underwent repeated phase I and II clinical trials. Given diminished efficacy and absent standard clinical trial progression in non-FDA-approved indications, we recommend an increased consideration of the potential benefits and adverse events patients face in clinical trials seeking to expand the indication of lenvatinib beyond its currently approved FDA indications.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11523-024-01040-5.

Declarations

Funding No external funding was used in the preparation of this manuscript.

Conflicts of Interest/Competing Interests Vinay Prasad reports research funding from Arnold Ventures; royalties from Johns Hopkins Press, Medscape, and MedPage; honoraria from GrandRounds/lectures from universities, medical centers, non-profits, professional societies, YouTube, and Substack; consulting from UnitedHealthcare and OptumRX; and speaking fees from Evicore. Plenary Session podcast has Patreon backers. Matt Vassar reports receipt of funding from the National Institute on Drug Abuse, the National Institute on Alcohol Abuse and Alcoholism, the US Office of Research Integrity, Oklahoma Center for Advancement of Science and Technology, and internal grants from Oklahoma State University Center for Health Sciences, all outside of the present work. Patrick Crotty, Karim Kari, Griffin K. Hughes, Chase Ladd, Ryan McIntire, Brooke Gardner, Andriana M. Peña, Sydney Ferrell, Jordan Tuia, Jacob Cohn, and Alyson Haslam declare that they have no conflicts of interest that might be relevant to the contents of this manuscript.

Ethics Approval The OSU-CHS Institutional Review Board (IRB) has determined that this project does not qualify as human subject research as defined in 45 CFR 46.102(d) and (f) and is not subject to further oversight by the OSU-CHS IRB.

Data Availability The datasets generated and/or analyzed during the current study are available in the OSF repository (https://osf.io/vdr68/?view_only=78fb8b6f2e3e440aa71423b40ec8dd46).

Author Contributions The work reported in the paper has been performed by the authors, unless clearly specified in the text. Each of the author's contributions have been reported in accordance with CRediT: PC: Conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing - original draft, writing - review & editing KK: conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing - original draft, writing - review & editing GH: conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing - original draft, writing - review & editing CL: conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing - original draft, writing - review & editing RM: data curation, investigation, methodology, project administration, supervision, validation, writing - original draft BG: data curation, investigation, methodology, project administration, supervision, validation, writing - original draft AP: data curation, investigation, methodology, project administration, supervision, validation, writing - original draft SF: data curation, resources, software, methodology JT: data curation, resources, software, visualizatio JC: data curation, resources, software, methodology AH: conceptualization, formal analysis, methodology, project administration, resources, validation, supervision, writing - review & editing VP: conceptualization, formal analysis, investigation, methodology, project administration, resources, supervision, validation, writing - original draft, writing - review & editing MV: conceptualization, investigation, methodology, project administration, resources, supervision, validation, writing - original draft, writing - review & editing

References

1. Kaitin KI. Deconstructing the drug development process: the new face of innovation. Clin Pharmacol Ther. 2010;87(3):356–61.
2. DiMasi JA, Grabowski HG, Hansen RW. Innovation in the pharmaceutical industry: new estimates of R&D costs. J Health Econ. 2016;47:20–33.
3. Lu Y, Dai Z, Chang F, et al. Whether and how disutilities of adverse events were used in the economic evaluation of drug therapy for cancer treatment. Pharmacoeconomics. 2023;41(3):295–306.
4. Wouters OJ, McKee M, Luyten J. Estimated research and development investment needed to bring a new medicine to market, 2009–2018. JAMA. 2020;323(9):844–53.
5. Hay M, Thomas DW, Craighead JL, Economides C, Rosenthal J. Clinical development success rates for investigational drugs. Nat Biotechnol. 2014;32(1):40–51.
6. Chhikara BS, Parang K. Global Cancer Statistics 2022: the trends projection analysis. Chem Biol Lett. 2023;10(1):451–451.
7. Carlisle B, Demko N, Freeman G, et al. Benefit, risk, and outcomes in drug development: a systematic review of sunitinib. J Natl Cancer Inst. 2016;108(1):djv292. https://doi.org/10.1093/jnci/djv292.
8. Carlisle BG, Zheng T, Kimmelman J. Imatinib and the long tail of targeted drug development. Nat Rev Clin Oncol. 2020;17(1):1–3.
9. Kimmelman J, Carlisle B, Gönen M. Drug development at the portfolio level is important for policy, care decisions and human protections. JAMA. 2017;318(11):1003–4.
10. Ouimet C, Bouche G, Kimmelman J. Diminishing clinical impact for post-approval cancer clinical trials: a retrospective cohort study. PLoS ONE. 2022;17(9): e0274115.
11. LENVIMA (lenvatinib) Label. https://www.accessdata.fda.gov/drugsatfda_docs/label/2021/206947s021lbl.pdf. Accessed 10 June 2023.
12. Center for Drug Evaluation, Research. Novel drug approvals for 2015. US Food and Drug Administration. Available at: https://www.fda.gov/drugs/new-drugs-fda-cders-new-molecular-entities-and-new-therapeutic-biological-products/novel-drug-approvals-2015. Accessed 17 May 2023.
13. Fogel DB. Factors associated with clinical trials that fail and opportunities for improving the likelihood of success: A review. Contemp Clin Trials Commun. 2018;11:156–64.
14. Foster ED, Deardorff A. Open Science Framework (OSF). J Med Libr Assoc. 2017;105(2):203.
15. Hughes G, McIntire R, Hightower B, et al. 06/23 June projects: Drug landscapes. Published online May 31, 2023. Available at:

https://osf.io/vdr68/?view_only=78fb8b6f2e3e440aa71423b40ec8dd46. Accessed 23 Jun 2023.
16. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan - a web and mobile app for systematic reviews. Syst Rev. 2016;5(1):210.
17. Araujo D, Greystoke A, Bates S, et al. Oncology phase I trial design and conduct: time for a change—MDICT Guidelines 2022. Ann Oncol. 2023;34(1):48–60.
18. Mattina J, Carlisle B, Hachem Y, Fergusson D, Kimmelman J. Inefficiencies and patient burdens in the development of the targeted cancer drug sorafenib: a systematic review. PLoS Biol. 2017;15(2): e2000487.
19. Hey SP, Heilig CM, Weijer C. Accumulating Evidence and Research Organization (AERO) model: a new tool for representing, analyzing, and planning a translational research program. Trials. 2013;14:159.
20. Makker V, Colombo N, Casado Herráez A, et al. Lenvatinib plus pembrolizumab for advanced endometrial cancer. N Engl J Med. 2022;386(5):437–48.
21. Motzer R, Alekseev B, Rha SY, et al. Lenvatinib plus pembrolizumab or everolimus for advanced renal cell carcinoma. N Engl J Med. 2021;384(14):1289–300.
22. Vergote I, Powell MA, Teneriello MG, et al. Second-line lenvatinib in patients with recurrent endometrial cancer. Gynecol Oncol. 2020;156(3):575–82.
23. Curtin F, Schulz P. Assessing the benefit: risk ratio of a drug–randomized and naturalistic evidence. Dialogues Clin Neurosci. 2011;13(2):183–90.
24. Rittberg R, Czaykowski P, Niraula S. Feasibility of randomized controlled trials for cancer drugs approved by the food and drug administration based on single-arm studies. JNCI Cancer Spectr. 2021;5(4):pkab061. https://doi.org/10.1093/jncics/pkab061.
25. Hutchinson N, Bicer S, Feldhake E, et al. Probability of regulatory approval over time: a cohort study of cancer therapies. JCO Oncol Pract. 2024;20(2):247–53.
26. Chapter 5: Collecting data. Available at: https://training.cochrane.org/handbook/current/chapter-05. Accessed 15 Jun 2023.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Patrick Crotty[1] · Karim Kari[1] · Griffin K. Hughes[1] · Chase Ladd[1] · Ryan McIntire[1] · Brooke Gardner[1] · Andriana M. Peña[1] · Sydney Ferrell[5] · Jordan Tuia[4] · Jacob Cohn[3] · Alyson Haslam[4] · Vinay Prasad[4] · Matt Vassar[1,2]

✉ Patrick Crotty
patrick.crottyosu@gmail.com

Karim Kari
karimkari.vassar@gmail.com

Griffin K. Hughes
griffinhughesresearch@gmail.com

Chase Ladd
cladd44@gmail.com

Ryan McIntire
ryanmcvt9@gmail.com

Brooke Gardner
Brooke.Hightower1.25@gmail.com

Andriana M. Peña
andriana.pena@okstate.edu

Sydney Ferrell
sydneyc.ferrell@gmail.com

Jordan Tuia
jordan.tuia@ucsf.edu

Jacob Cohn
taylorlynncohn@gmail.com

Alyson Haslam
alyson.haslam@ucsf.edu

Vinay Prasad
vinayak.prasad@ucsf.edu

Matt Vassar
matt.vassar@okstate.edu

[1] Office of Medical Student Research, Oklahoma State University Center for Health Sciences, 1111 W 17th St., Tulsa, OK 74107, USA

[2] Department of Psychiatry and Behavioral Sciences, Oklahoma State University Center for Health Sciences, Tulsa, OK, USA

[3] Department of Internal Medicine, Oklahoma State University Center for Health Sciences, Tulsa, OK, USA

[4] Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA, USA

[5] Department of Internal Medicine, University of Vermont Medical Center, Burlington, VT, USA