Epidemiology and characteristics of clinical trials supporting US FDA approval of novel cancer drugs

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Aviv Ladanie

aus Glarus Süd, Schweiz

Basel, 2018

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von Prof. Dr. Jürg Utzinger (Fakultätsverantwortlicher), PD Dr. Lars G Hemkens (Dissertationsleiter) und Prof. Dr. André Knottnerus (Korreferent).

Basel, den 18. September 2018

Prof. Dr. Martin Spiess

Dekan

Table of contents

Acknowledgments

This thesis and the included manuscripts would not exist without the help of numerous people and organizations who directly or indirectly supported either me or my PhD project in various ways over the years. I would like to acknowledge the major role of Lars G Hemkens as PhD thesis supervisor and principal investigator of the "Comparative Effectiveness of Innovative Treatments for Cancer" (CEIT-Cancer) project. I thank him for giving me a high level of autonomy in managing the first part of the CEIT-Cancer project and his trust in me as a competent scientist. I would like to give credit to Heiner C Bucher who allowed me to be part of his institute and who acted as the initial thesis supervisor. I would like to acknowledge the importance of the thesis committee, Jürg Utzinger and André Knottnerus, for their time and efforts to ensure that my PhD progressed continuously over the years. I am also greatly indebted to Diana Grauwiler, Anja Schreier, Anneke Germeraad, and Christine Mensch for their unconditional support and kindness whenever I needed their help. My big thanks go to Benjamin Kasenda, who not only supported the CEIT-Cancer project from the beginning on with his disease-specific expertise but also established the contact with other oncologists. The latter was of immense importance for the completion of the project in a reasonable amount of time. My special thanks go to Arnav Agarwal, Matthias Briel, Hannah Ewald, John PA Ioannidis, Juan Martin-Liberal, Florian Naudet, Tiago V Pereira, Thomas Schmid, Francesco Sclafani, and Benjamin Speich for their strenuous efforts to retrieve information from a messy data source. I thank the Swiss Cancer League for making my PhD project possible by funding the CEIT-Cancer project. I also highly appreciated the opportunities to attend courses not offered by Swiss universities and to present my work at international conferences. All this would not have been possible without the financial support of the PhD Program Health Sciences (PPHS), the Swiss School of Public Health (SSPH+), and the Travel funds of the University of Basel.

Finally, I thank my family and friends for their continuous support throughout my PhD. I am particularly grateful to my wife, Elena, for her patience, understanding, and support that I needed to make this thesis possible.

Summary

The US Food and Drug Administration (FDA) approves novel drugs that appear to be effective for their intended uses and whose benefits outweigh their risks. The legal standards for drug approval are widely understood to require evidence of efficacy from two or more clinical trials that independently demonstrate statistically significant effects in favor of the experimental drug and on endpoints that reflect a clinical benefit to patients. However, the FDA has the authority to be flexible in applying the approval standards, particularly in the case of drugs that are intended to treat serious medical conditions. This regulatory flexibility, however, is frequently and repeatedly criticized for putting patients at risk.

The two overall aims of this thesis were a) to develop a guide on how to access, retrieve, and use of clinical trial data that supported FDA approval of novel treatments published by the agency itself; and b) to describe characteristics and extent of clinical trial evidence that supported FDA approval of novel drugs for cancer indications between 2000 and 2016. These aims were addressed in three manuscripts, and we describe the methods used to retrieve and manage the clinical trial information in a fourth manuscript.

Manuscript 1: How to use FDA drug approval documents for evidence syntheses

The FDA publishes information about the clinical trial evidence that supported approval of novel drugs and therapeutic biologics in the drugs@FDA database in form of “drug approval packages”. Information in the main document extends over hundreds of pages and is structured in a way that may be unintuitive for researchers. Although the value of this source of potentially unpublished clinical trial information is undisputed, its use in evidence syntheses of drug interventions remains limited. Based on our experience in using the drugs@FDA database and drug approval packages, we provide step-by-step instructions on how to navigate through the database as well as how to access, efficiently find and retrieve, and use the clinical trial information. Our guide may promote better use of this information,

which may improve the completeness and validity of future evidence syntheses of drug interventions.

Manuscript 2: The Comparative Effectiveness of Innovative Treatments for Cancer (CEIT-Cancer) project: rationale and design of the database and the collection of evidence available at approval of novel drugs

We describe the rationale and efforts made to identify all clinical trials that supported FDA drug approval between 2000 and 2016 for the treatment of cancer and to retrieve pertinent information about trial design and treatment effects on overall survival, progression-free survival, and objective tumor response. Most data retrieval steps were conducted by two data reviewers (who worked independently and who were guided by an instruction manual) to reduce random errors that would affect the quality of the collected data. The study design can be applied in the future for projects with similar scopes, and the collected data will be used in the future for numerous research projects.

Manuscript 3: Clinical trial evidence supporting US FDA approval of novel cancer therapies between 2000 and 2016

Using the data collected in the CEIT-Cancer project, we analyzed the 127 clinical trials that supported FDA approval of 92 novel drugs for the treatment of 100 cancer indications. The median number of enrolled patients was 193 (interquartile range [IQR]: 106, 448). Fifty-one percent (51%) were randomized controlled, and 75% were open-label. The hazard ratio (HR) for the pooled average treatment effect on overall survival was 0.77 (95% confidence interval [CI]: 0.73, 0.81; I-squared [I^2] = 47%), and HR = 0.52 (95% CI: 0.47, 0.57; I^2 = 88%) for progression-free survival. The odds ratio for objective tumor response was 3.58 (95% CI: 2.77, 4.62; I^2 = 87%). The median absolute survival gain was 2.40 months (IQR: 1.25, 3.89). These findings indicate that novel cancer treatments are supported by trials with design features that have the potential to threaten the validity of the findings and that the overall absolute survival difference is small.

Manuscript 4: Corroborating characteristics of single pivotal trial evidence supporting FDA approval of novel cancer therapies

For experimental new drugs intended to treat serious conditions, the FDA may grant marketing approvals based on evidence from a single trial alone (instead of two or more) if certain trial characteristics are met. The presence of one or more of these trial characteristics defined by the FDA may increase the FDA's confidence in the validity of a single clinical trial and may therefore provide corroborating evidence of efficacy. Our results show that 36 out of 100 approvals of novel cancer treatments were based on evidence from a single trial alone. Sixty-four percent (64%) were large and multicentric trials; 64% had consistent effects across different subgroups; 42% had consistent effects across endpoints; and 33% had very low p-values for the primary endpoint. Overall, 92% of clinical trials that supported FDA approval alone fulfilled at least one of the corroborating characteristics. Whether the presence of one or more of these corroborating characteristics indeed provides a safeguard against threats to the validity of trials remains to be answered.

The background information provided in the manuscripts and this thesis improve the understanding of the regulatory considerations that are made to bring novel cancer treatments to market. The results of the analyses provide insight into the characteristics of the clinical trial evidence and the number of clinical trials that supported drug approval in the fields of oncology and malignant hematology.

Important acronyms

BLA	Biologics License Application
CEIT-Cancer	Comparative Effectiveness of Innovative Treatments for Cancer
CI	Confidence interval
EMA	European Medicines Agency
FDA	Food and Drug Administration
HR	Hazard Ratio
IND	Investigational New Drug
IQR	Interquartile Range
NDA	New Drug Application
OR	Odds Ratio
OS	Overall Survival
PFS	Progression-Free Survival
RCT	Randomized Controlled Trial
SAT	Single-Arm Trial
TR	Tumor Response

1. General introduction

The Food and Drug Administration (FDA) is the authority in the United States (US) that regulates human medical products, including drugs and therapeutic biologics (1). The FDA's regulatory oversight begins when a drug research and development organization - usually a for-profit biopharmaceutical company - with an experimental drug or therapeutic biologic decides to enter the clinical development stage and to examine its efficacy and safety in humans. Before initiation of the first-in-human trial, the biopharmaceutical company must obtain permission to conduct experiments with human subjects by submitting an Investigational New Drug (IND) application to the FDA (2). An IND contains the full protocol of clinical trials to be conducted on US territories, and records are continuously updated to reflect changes to the clinical development program such as protocol amendments (3). The purpose of an IND is twofold: first, to ensure the safety and rights of human subjects enrolled in clinical trials; and second, to ensure that phase 2 and 3 trials that are expected to demonstrate drug efficacy are scientifically sound and meet FDA standards to support drug approval (details are elaborated below) (4). Clinical trials designed to demonstrate drug efficacy and to meet regulatory standards are referred to by different names, such as pivotal trials, phase 3 trials, or registration trials.

An IND holder may apply for regulatory permission to commercialize an investigational new drug or therapeutic biologic for a well-defined clinical use if efficacy has been sufficiently established in pivotal trials. The formal process to apply to the FDA for marketing authorization is via a New Drug Application (NDA) for chemically synthesized drugs or a therapeutic Biologics License Application (BLA) for biotechnology-derived products. An NDA or BLA dossier contains all the information collected during the IND stage (5), including detailed information on the composition of the drug or therapeutic biologic product, patient-level data from clinical trials of the product, and findings from statistical analyses of the clinical trial data (6). A multidisciplinary team of FDA employees scrutinizes the content of an application dossier for its scientific validity and compliance with regulatory standards. Each discipline involved in this process documents its findings and conclusions in the form of a review, which eventually

informs the FDA's decision of whether to approve the application. The most critical question to be answered for the FDA's approval decision is whether the benefits achieved with the drug or therapeutic product (hereafter collectively referred to as "drug") in its targeted medical condition outweigh the known and potential harms to future patients ("positive benefit-risk profile") (7).

1.1. FDA standards for drug approval

According to legal requirements, determination of the benefit-risk profile has to be made based on data that provides "substantial evidence" from "adequate and well-controlled investigations", based on which it could be concluded "that the drug will have the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the labeling or proposed labeling thereof" (8). As the final rulemaking authority, the FDA defined the key design and analysis features of an adequate and well-controlled investigation as follows (9):

- "The study uses a design that permits a valid comparison with a control to provide a quantitative assessment of drug effect", with either a concurrent placebo, dose-comparison, no treatment, or active control, or a historical control,
- "The method of assigning patients to treatment and control groups minimizes bias and is intended to assure comparability of the groups" and furthermore, "in a concurrently controlled study, assignment is by randomization",
- "Adequate measures are taken to minimize bias on the part of the subjects, observers, and analysts of the data […] such as blinding".

More granular information about drug approval standards and their interpretation as well as FDA recommendations for drug development can be found in "Guidance for Industry" documents and peer-reviewed articles published by FDA staff. With regards to the substantial evidence requirement, the FDA clarified that the use of the plural in "adequate and well-controlled investigations" is interpreted as the legislative intent to require data from at least two clinical trials (10). This interpretation is supported by the fundamental scientific principle of replication by which the credibility of findings from

a single study can be strengthened by validation in a second, independent experiment. Consequently, the FDA generally requires biopharmaceutical companies to conduct at least two clinical trials. In 1997, however, US Congress amended the legal requirement to allow approvals based on positive evidence from only a single pivotal trial under exceptional circumstances and at the discretion of the FDA (10). Additional elements of the substantial evidence requirement include a demonstration of statistically significant effects at the five percent level (8) on clinically meaningful endpoints, which encompasses both endpoints that provide direct evidence of clinical benefit (i.e., "clinical endpoints", defined by the FDA as those that "directly measure what matters most to people—whether they feel or function better, or live longer") (11) as well as (in view of the FDA) sufficiently validated surrogate endpoints (12). The latter type of endpoint measures a proxy for clinical benefit which is correlated (for a specific disease setting and drug class) with a clinical endpoint (13), such as Hemoglobin A1c [HbA1c] for diabetic long-term complications.

1.2. Regulatory flexibility and expedited programs for serious conditions

The complete lack of effective treatment since the emergence of Human Immunodeficiency Virus/Acquired Immune Deficiency Syndrome (HIV/AIDS) in the early 1980s until 1987 brought the FDA into the spotlight of advocates of HIV/AIDS patients who equated FDA regulations and policies for drug development and approval with administrative bureaucracy (14, 15). They urged the agency to enable faster availability of HIV/AIDS drugs by providing access as soon as an investigational new drug is shown to be safe and demonstrates early evidence of efficacy.

Confronted with this demand, the FDA came to the recognition that a single interpretation of the standards for drug approval does not suit the needs of all patients (16). Of crucial importance was the understanding that some patients living with severe medical conditions and with an urgent need for new treatment options are willing to take experimental drugs that have demonstrated early but not definitive evidence of efficacy (17, 18). In the following years, the FDA and US Congress introduced several policy and legislative reforms aiming to speed up development, evaluation, and approval of new

drugs. They are all based on a principle proposed by the FDA in 1988 (19) that gave the FDA the authority to exert flexibility in applying the drug approval standards, mainly by using its scientific judgment to determine the amount and type of evidence needed for approval of novel drugs (20). This regulatory flexibility is only applicable to drugs that are intended for the treatment of serious conditions (21).

According to FDA definitions, the term "serious conditions" encompasses diseases and medical conditions that are associated with irreversible morbidity and substantial impact on day-to-day functioning, as well as life-threatening conditions that have either a high likelihood of death (if left untreated) or are associated with potentially fatal outcomes (19, 22). Examples include HIV/AIDS, Alzheimer's disease, heart failure, or cancer (23), but also many rare diseases (19). Rare diseases are defined in the US as medical conditions that affect less than 200,000 individuals in the US and are often referred to as "orphan diseases" (and the drugs intended for rare disease patients as orphan drugs or as drugs with an orphan status) (24).

1.2.1. Procedures to expedite drug development, review, and approval

The FDA formulated the principle of regulatory flexibility in four expedited programs that are intended to speed up drug development, review, and approval of novel drugs for serious conditions (19). The first major program was implemented in 1992 with the accelerated approval pathway as an alternative to traditional approval (25). Accelerated approval can be granted to drugs intended to treat serious conditions that have the potential to address an unmet medical need (19). In contrast to traditional marketing approval (requiring demonstration of substantial evidence of clinical benefit), accelerated approval is granted if evidence of efficacy is demonstrated on an unvalidated surrogate endpoint that is in view of the FDA "reasonably likely to predict clinical benefit". However, accelerated approval is a conditional marketing authorization, and the marketing authorization holder is obliged to verify the clinical benefit in a confirmatory trial after approval to convert the status from temporary to full marketing approval. Failure to do so could theoretically lead to the withdrawal of the accelerated

approved marketing authorization.

Other programs to expedite drug development and review were priority review introduced in 1992, fast track introduced in 1997, and breakthrough therapy introduced in 2012. Their qualifying criteria and regulatory actions are presented in Table 1. However, it is important to note that official standards for drug approval remain unaltered except for the accelerated approval program where approval is granted upon demonstration of efficacy based on unvalidated surrogate endpoints (that are in view of the regulators reasonably likely to predict clinical benefit) instead of clinical endpoints. The remaining programs aim to either expedite application review times (priority review) or both development and review times (fast track and breakthrough therapy).

TABLE 1: FDA PROGRAMS TO EXPEDITE DRUG DEVELOPMENT AND REVIEW FOR SERIOUS CONDITIONS

	Priority Review	Fast Track *	Breakthrough Therapy *
Enactment year	1992	1997	2012
Qualifying criteria for drugs	Assigned if it provides a significant improvement in safety or effectiveness.	Assigned if nonclinical or clinical data demonstrate the potential to address an unmet medical need.	Assigned if preliminary clinical evidence indicates that the drug may demonstrate substantial improvement on a clinically significant endpoint over available therapies.
Regulatory program features	Application review and final decision are made within six months of application submission (whereas it takes ten months under standard review) by directing additional resources.	• Increased interaction between the FDA and developing biopharmaceutical company to discuss the development plan. • Completed NDA or BLA sections are reviewed by the FDA before a biopharmaceutical company submits the final application (rolling review).	• Increased interaction between the FDA and biopharmaceutical company to discuss the design and conduct aspects of a drug development program. • Involvement of senior staff and management. • Completed NDA or BLA sections are reviewed by the FDA before a biopharmaceutical company submits the final application (rolling review).

* Designation can be revoked if accumulating evidence indicates that the qualifying criteria are no longer met. Adapted from (19).

1.3. Criticism and extent of regulatory flexibility

Parts of the medical research community view the application of regulatory flexibility and the use of programs aiming to expedite drug development, review times, and approval with great skepticism, particularly for cancer drugs (26-28). They criticize that the robustness of pivotal trial evidence may be compromised, implicitly because findings are not replicated in independent trials (i.e., a single positive finding may represent a chance or biased result) and because trials are smaller (which for example may lead to more imprecise estimates of efficacy and safety). They also criticize the use of surrogate endpoints (with questionable relevance to patients) and argue that the internal validity of findings from non-randomized, historically controlled trials may be threatened by various biases. Together, this may increase uncertainty for decision-making about the benefits and harms of new drugs for various stakeholders in the healthcare system such as the FDA (who makes approval decisions), health insurers (who make coverage decisions), as well as physicians and patients (who make treatment decisions), which may ultimately lead to approval, reimbursement, and use of drugs that are ineffective, unsafe, or both.

Several studies analyzed the extent of regulatory flexibility by examining the design characteristics and number of pivotal trials that support approval of novel drugs. The landmark article by Downing et al. published in 2014 (29) mainly revealed large variations in trial design features between various therapeutic areas (such as infectious diseases or psychiatry). Compared to all therapeutic areas combined, pivotal trials supporting FDA approval of drugs for the treatment of cancer diseases were less likely to be randomized (47% versus 89%), double-blinded (27% versus 80%), have a comparator group (13% versus 53%), and measure clinical endpoints as a primary endpoint (16% versus 29%). The median number of enrolled patients was also lower (266 versus 446). In addition, 80% of all approvals were supported by data from only a single trial (compared to 37% across all therapeutic areas). Kesselheim et al. (30) analyzed the association between orphan status and trial design characteristics. They estimated that pivotal trials of drugs for rare cancers were less likely to be randomized (30% versus 80%) or double-blind (4% versus 33%). They were also less likely to measure overall survival as the

primary endpoint (8% versus 27%). In addition, they had a lower median number of enrolled patients (96 versus 290). Finally, concerning the number of pivotal trials, Downing et al. (29), Sridhara et al. (31), Martell et al. (32), Morant et al. (33), Gentry (34), and Tibau et al. (35) estimated that 80% to 85% of cancer drug approvals are based on evidence of treatment efficacy from only a single pivotal trial.

2. Thesis overview

2.1. Rationale and context

Although there was a tangible body of evidence on the extent of regulatory flexibility for the approval of novel drugs already existing at the time of initiation of this thesis in 2015, the knowledge specifically on cancer drugs was fragmented across the numerous publications, and cross-comparisons were complicated by methodological differences. In addition, Downing et al. (29) missed the opportunity to evaluate the influence of orphan status or accelerated approval specifically for the different therapeutic areas. Kesselheim et al. (30) examined the impact of orphan status on trial design features but ignored the accelerated approval program. Furthermore, the analysis was based on a small sample of only 27 cancer drugs. Finally, the proportion of single pivotal trials was derived in an oversimplified manner by merely determining the number of trials discussed in regulatory documents (29, 31-35).

The present thesis is embedded in a larger project that strives to describe the collective scientific knowledge on the efficacy of novel drugs and its generation over time, the "Comparative Effectiveness of Innovative Treatments" (CEIT) project, which started with a focus on cancer (CEIT-Cancer). As the first part of this overarching project, and in addition to the aims and objectives listed below, this thesis laid the foundation for subsequent CEIT-Cancer subprojects by (a) generating knowledge of the FDA regulatory framework for drug development, review, and approval, (b) identifying the drugs of interest, and (c) retrieving the efficacy evidence generated in the clinical development phase (also known as the "pre-marketing" period of drug development) of these drugs in a systematic and reproducible manner.

Two factors drove our motivation to use information published and curated by the FDA over the traditional peer-reviewed medical literature. First, FDA drug approval reviews may provide more comprehensive (36) and independent (37) insight into the clinical trial evidence generated by biopharmaceutical companies during the clinical development phase of new drugs. This information is publicly available in the drugs@FDA online database (38). Second, numerous attempts were made in

the past to familiarize secondary users of clinical trial data such as authors of evidence syntheses of drug interventions with this comprehensive and detailed source of clinical trial information (with the overall goal to increase its use for research purposes). But despite these attempts, it is still underused, presumably because researchers are unfamiliar with the regulatory context or because they are discouraged by the wealth of information. The first factor allowed us to build a comprehensive inventory of the pre-marketing clinical trials of novel cancer drugs using publicly available information. We regarded the second factor as an opportunity to provide other investigators with our experience in using FDA drug approval reviews for research purposes and to help filling a knowledge gap (39).

2.2. Aims and objectives

The thesis aims to develop procedures that facilitate access, retrieval, and use of clinical trial data provided by the FDA and aims to describe the clinical trial evidence that supported FDA marketing approval of novel cancer drugs between 2000 and 2016. The rationale and context are outlined earlier in this chapter.

Procedures to facilitate access, retrieval, and use of clinical trial data from drugs@FDA for evidence syntheses of drug interventions:

- Objective I: to update and expand an outdated guide on how to access and retrieve drug approval packages created and released by the FDA;
- Objective II: to develop standard operating procedures on how to locate, retrieve, and use clinical trial information provided in FDA drug approval documents.

Characteristics and extent of clinical trial evidence supporting FDA approval of novel cancer drugs between 2000 and 2016:

- Objective III: to describe the clinical trial evidence that supported FDA marketing approval
- Objective IV: to examine the association of disease and regulatory characteristics with features and extent of the pivotal trial evidence;
- Objective V: to determine how frequently FDA approvals are granted based on data from single

clinical trials and to describe the circumstances.

2.3. Structure

These aims and objectives are addressed in three individual manuscripts provided in chapters 3, 5, and 6. Chapter 3 presents findings of objectives I and II, chapter 5 of objectives III and IV, and chapter 6 of objective V. In addition, the fourth manuscript in chapter 4 describes the substantial efforts made in the CEIT-Cancer project to identify, select, retrieve, and manage the data used for answering objectives III, IV, and V. It can be considered in part to be the detailed methods section of the manuscripts presented in chapters 5 and 6.

3. How to use FDA drug approval documents for evidence syntheses

Aviv Ladanie [1, 2], Hannah Ewald [1, 2, 3], Benjamin Kasenda [1, 4], Lars G Hemkens [1]

[1] Basel Institute for Clinical Epidemiology and Biostatistics, Department of Clinical Research, University Hospital and University of Basel, Basel, Switzerland
[2] Swiss Tropical and Public Health Institute (Swiss TPH), Basel, Switzerland
[3] University Medical Library, University of Basel, Basel, Switzerland
[4] Departments of Medical Oncology and Haematology, University Hospital and University of Basel, Basel, Switzerland

Status: The final version of this manuscript is published in: the BMJ 2018;362:k2815. https://doi.org/10.1136/bmj.k2815

Evidence syntheses may benefit from using aggregated clinical trial information in approval documents published online by the US Food and Drug Administration (FDA). Here we provide practical guidance on how to access and use this source of information for evidence syntheses on treatment effects of drugs and therapeutic biologics.

SUMMARY BOX

- There is compelling evidence that published trial information is selectively reported and that results not showing favourable effects of the tested treatments often remain unpublished.
- Clinical trial information published by regulatory authorities such as the FDA may help to address and reduce such reporting biases.
- FDA approval documents are very long and do not follow the typical structure of medical journal articles which may discourage reviewers to use them for evidence syntheses
- Our practical guidance on how to efficiently identify and use approval documents to find the relevant information may help promoting the use of this valuable data source for evidence syntheses.

Publicly accessible US FDA approval documents allow gaining important insights into reporting biases in articles published in peer-reviewed medical journals (1-4), but this data can also be used for other purposes: to directly minimize the impact of such biases on the results and conclusions in evidence syntheses (5), to obtain information not disclosed in published clinical trials reports (6), or to identify unpublished clinical trials to increase precision of effect estimates (7).

For example, almost 20 years ago, Man-Son-Hing and colleagues (5) showed that incorporating unpublished trials into a meta-analysis on quinine for nocturnal leg cramps substantially reduced the estimated efficacy. The bias occurred because almost all published trials had larger effects than the unpublished studies. Similarly, Turner and colleagues found that 22 of 71 (31%) trials discussed in FDA approval documents of 12 antidepressants were not published, and that publication was closely

associated with results favouring the experimental drugs (1). Hart and colleagues showed that updating meta-analyses with unpublished trial data from FDA approval documents changed drug efficacy estimates in 38 of 41 cases (93%) towards both lower and higher efficacy (4). Rising and colleagues not only revealed that approval trials are frequently unpublished, in particular when they suggest unfavourable outcomes for the experimental intervention, but also that the published information is incomplete because results were omitted from the papers (2). MacLean and colleagues, on the other hand, demonstrated that incorporating unpublished trial data in meta-analyses does not necessarily change treatment effect estimates (7). McDonagh and colleagues show various examples where information from FDA documents would alter conclusions of drug effectiveness reviews, not only by identifying unpublished studies, but also by providing unpublished information on benefits and harms found in published trials and by reporting results of independent analyses conducted by the FDA (6).

Overall, such examples illustrate that unpublished FDA trial data has the potential to change the results of evidence syntheses and can provide useful information that would otherwise be unavailable (6, 8). It may help to better understand the strengths and weaknesses of a trial that only a regulator could discover, given the regulator's access to the original study data and the original trial protocols (drug developers must submit the trial protocols before they initiate the approval trials (9)).

However, regulatory data is rarely used in evidence syntheses. A survey estimated that only 24 of 794 Cochrane reviewers (3%) who had searched for unpublished clinical trials had gathered information from health authorities (10). The survey authors hypothesise that "some authors might not be aware of the amount of accessible data at regulatory agencies". Others emphasise that FDA approval documents are difficult to access and navigate (11-14).

Attempts to promote the adoption of regulatory data as a viable source of trial information have been made recently, including a guide on where and how to retrieve FDA approval documents (13), a description of the content in FDA and European Medicines Agency (EMA) documents (15), and the dissemination of FDA approval documents in a more accessible format via the recently launched OpenTrialsFDA platform (16). The sheer amount of information encountered in FDA approval

documents, usually hundreds of pages not following the typical structure of medical journal articles, may indeed confuse and discourage reviewers to use them for evidence syntheses.

Guidance could be provided by either the publisher of these documents (the FDA), leading organizations advocating for systematic reviews of healthcare interventions (particularly the Cochrane Collaboration), or reviewers with experience in using FDA approval documents.

To our knowledge, the FDA has not published any description of the content and structure of approval documents, in contrast to the EMA and the Australian Therapeutic Goods Administration (TGA) (17). Neither does the Cochrane Handbook provide advice in this direction, and – according to the changelog (18) – the updated version (5.2, June 2017) will not be addressing this issue either.

We use FDA approval documents in meta-epidemiological projects, including an ongoing analysis of 92 anticancer agents approved by the FDA between 2000 and 2016. Here, we would like to share our knowledge and describe how we navigate such documents efficiently. We also indicate where one can expect to find information typically relevant for evidence syntheses.

The target users of this guide are mainly authors of evidence syntheses who intend to collect and synthesise evidence on a given topic in a systematic and transparent manner, such as in systematic reviews or meta-analyses. This guidance can assist in identifying potentially unpublished drug trials, obtaining additional information that is unavailable from published clinical trial reports, or for cross-checking information reported in journal articles. However, it does not cover the subsequent indispensable steps of evidence synthesis including the quality assessment, for which further detailed guidance would be needed.

The FDA drug approval package

To obtain marketing authorisation for newly developed drugs and therapeutic biologic products (herein referred to as “drugs”), companies have to submit a New Drug Application (NDA) or Biologics License Application (BLA) to the FDA and provide information about the drug’s quality, safety, and efficacy (19).

The FDA reviews this information and - for drugs that are ultimately approved for marketing – publishes these reviews (albeit in a form redacted of some information) online in the drugs@FDA database as PDF-documents (20). Documents pertaining to a single approved product are organized in "approval packages". There are a number of guidance (21) and policy documents (22) which provide a deeper understanding of the FDA processes and evolving procedures. The review process is addressed by a "Good Review Practice" document within FDA's Center for Drug Evaluation and Research (CDER) "that discusses any aspect related to the process, format, content and/or management of a product review" (23).

Approval packages are available for prescription and over-the-counter drugs as well as "drug-like" agents (such as therapeutic biologics that encompass antibodies, cytokines, growth factors, enzymes, immunomodulators and others). For a detailed list of what is included in the drugs@FDA database see reference (24). Approval documents are only sporadically available for drugs approved prior to 1997, and are rarely available for supplemental indications (i.e., indications approved once a drug received its initial marketing authorization). Approval documents for devices and non-therapeutic biologics (such as vaccines, blood and blood products) are not regulated by the CDER and are not addressed by this guidance.

Finding and accessing approval packages

We provide guidance on determining whether FDA approval packages are available for a given drug of interest and how to find the corresponding FDA approval package in the supplemental material ("Part 1: How to access FDA approval packages").

Document types in approval packages

We are aware of about 20 different document types included in FDA approval packages (Box 1). The medical review is typically most relevant for evidence syntheses. The medical reviews (sometimes referred to as "Clinical Review") usually contain sufficient information for identifying and selecting

pertinent trials as well as main information about clinical trial characteristics, statistical analyses and results. However, some important details may only be revealed after thorough perusal of the entire approval package. Since 2016, the medical review document of recently approved drugs is often merged with other document types in a single "Multi-discipline Review" document. Its content can now be found under the table of contents heading "Statistical and Clinical Evaluation". If the medical review document is missing, incomplete, or illegible, or when more in-depth analysis is required, the other documents available in the FDA approval package may provide further information (see below).

Medical review structure

The medical review document structure has evolved over the years, but we identified three general document structures (Figure 1) used since approximately 2004. For older drugs and biologicals, there is no such consistent document structure.

The actual document structure may deviate slightly from these general structures, for example by section heading (e.g. section "Review of Efficacy" may be titled "Efficacy Evaluation" or "Integrated Review of Efficacy"), grouping of information (e.g. section "Clinical Pharmacology" may be a subsection of the "Significant Issues from Other Review Disciplines" section, or it may exist as a standalone section), or sequence of sections.

We generally suggest that reviewers first try to locate the table of contents (where available, some older reviews may have none) and make themselves familiar with the document structure. Sometimes there may be two table of contents in a medical review document. This may indicate that the original application for marketing authorization was declined by the FDA and that the agency re-evaluated the drug in a second review. Triptorelin/TRELSTAR® (25) is such an example (with table of contents on pages 2 and 41). In this situation, both medical reviews should be scrutinized because the data may differ between the two. Whether to use the superseded or the updated version can only be decided in the context of the research question. An explanation why triptorelin was only approved after a second review cycle can be found in the regulatory history, usually described in sections titled "Introduction",

"Background" or similar (Figure 1, greyed out as usually less relevant for evidence syntheses).

Data collection

Where to find relevant clinical trials and their characteristics

Trials submitted to the FDA to support approval are presented and discussed in detail in the pink highlighted sections in Figure 1. There is typically a tabular overview with brief information on individual trials such as the target population, interventions, comparators, outcomes, time frame, setting, and study design.

Many trials have multiple trial names or identifiers (for example C0743T09 and PHOENIX 2 in the case of ustekinumab/STELARA® (see medical review page 17 (26)), or CLEOPATRA, TOC4129g, or WO20698 in the case of pertuzumab/PERJETA® (see medical review page 30 (27)). Knowing all trial identifiers may facilitate locating the corresponding record on clinicaltrials.gov or identifying reports of trials published in journals. For example, the pivotal trial of ustekinumab/STELARA® can be found as "PHOENIX 2" on PubMed but only with detours on clinicaltrials.gov, while it can be quickly found with its identifier C0743T09 on clinicaltrials.gov but not in PubMed.

The pink highlighted sections are comparable with the methods section in medical journals in terms of content, with information on trial design (objectives; geographical distribution of sites), trial population (eligibility criteria), interventions, trial endpoints (definitions; outcome assessment), and statistical methods (sample size; calculation of effect estimates and details on interim, subgroup, and sensitivity analyses). Excerpts from trial protocols and a history of protocol amendments may also be found in these sections. They allow, for example, assessing pre-specification of endpoints or subgroups. For drugs approved before 2007, additional information about trial methods (and results) may be available in the Appendix (Figure 1).

Where to find trial results

Results for efficacy endpoints are reported in the green highlighted sections in Figure 1. The sections start with details on the trial population and enrolment, often including a flow diagram, site-specific enrolment information, and summary statistics of patient characteristics.

Not all trials presented in the pink highlighted sections (Figure 1) have results, either because they are of no relevance for drug approval, ongoing, or because they address a different indication. Such trials may still have been used to evaluate the drug's safety profile and to better understand the risk for adverse events. The methods and results of the safety review are presented in the yellow highlighted sections (Figure 1).

We provide a step-by-step instruction (including a working example) on where to find relevant clinical trials and their characteristics in the supplemental material.

Extraction of trial results

Numerous results for various endpoints from several pre-specified and exploratory analyses are presented in the results section of the medical review. The decision about which one to choose should be made depending on the research aims. We present some general examples of the various types of analyses reported in medical reviews to facilitate pre-specification of the extractions and analyses of interest (Box 2). We provide a step-by-step instruction (including a working example) on where to find trials results in the supplemental material.

Further and more detailed information

Additional information on statistical analyses and sample size calculation are often provided in the statistical review document which is also included in FDA approval packages. FDA guidance states that "applicants are expected to submit data of high quality and make it possible for the FDA to reproduce their results. In turn, FDA reviewers should provide adequate documentation so that the applicant or another data user could reproduce their independent findings" (28). The statistical review often

includes details of such re-analyses, for example "whether it is possible to reproduce the primary analysis dataset, and in particular the primary endpoint, from the original data source" (28).

Sometimes there are comments by the FDA medical or statistical reviewer which may be very informative for the assessment of the quality of evidence, for example to address potential risks of bias, or to discuss the adequacy of comparator interventions.

Most recently (March 2018), the FDA announced to publish more detailed information, such as Clinical Study Reports (CSR). This process is in a very early stage and it remains to be hoped that such promising evidence will be available for future drug approvals, but also from drugs approved in the past. There is currently the option to obtain such CSRs through Freedom of Information Act requests (29).

Challenges and potential solutions

The reporting quality in FDA approval documents has generally improved over time, but inconsistencies and contradictions across medical review sections and document types occur. Independent data extractions by two reviewers may be helpful to overcome this problem and to increase the reliability of the data. Recording on which document pages the extracted information was found may facilitate the consensus steps. A general problem of using the approval package is missing, inconsistent or selectively reported information. A specific problem is that some of the information in FDA reports may be redacted for various reasons (15).

Schroll and colleagues reported that in their systematic sample of drugs approved between 2011 and 2012 "crucial information about safety concerns and nonapproved indications were redacted in the FDA reports" (15). Therefore, utmost attention and careful evaluation on a case by case basis is required to assess potential biases resulting from incomplete information. In this regard, a close evaluation of the concerns of the various FDA reviewers in assessing efficacy and safety may provide valuable information. Drugs approved before 2007 are more prone to suboptimal reporting. Options to deal with missing information for outcome data include indirect calculation of effect estimates of time-to-event endpoints (e.g., arm-specific number of events, point estimates, p-values, Kaplan-Meier curves) (30,

31), and juxtaposition of the FDA approval document information with corresponding trial reports in published journal articles, trial registries (such as clinicaltrials.gov), or material from other approval agencies.

This guidance aims to reduce barriers to use this evidence by making access and navigation easier. However, numerous important questions remain about how to best integrate this valuable source of information into evidence syntheses. We touch only briefly upon crucial subsequent steps of evidence synthesis, for which further detailed guidance is needed. In particular, thorough assessment of the quality and potential biases of the information provided in approval documents is indispensable. Overall, we believe that despite such limitations, the consideration of approval documents can strengthen evidence syntheses of drug interventions.

Footnotes

Acknowledgement

Parts of this paper were presented at the Evidence Live conference in Oxford UK (June 21-22, 2017) under the title “Assessing drug intervention effects with published FDA approval summary documents: an experience report and practical guidance”. We thank two colleagues at the University Hospital and University of Basel (Dominik Glinz, Basel Institute for Clinical Epidemiology and Biostatistics) and (Christopher Traenka, Department of Neurology) for pilot testing the tutorial.

Provenance

We describe the structure and content of FDA approval documents based on our experience from various meta-research projects, including the ongoing CEIT-cancer project in which we evaluate the evidence base from pre-marketing clinical trials of 92 anticancer agents. In this project, we systematically acquire FDA approval packages, peruse approval documents and extract trial

characteristics as well as treatment effect estimates. In addition, we pilot tested the applicability of the manual for a number of drugs approved for neurological, cardiovascular, psychiatric, endocrinologic, and rheumatologic disorders and can confirm its validity across medical specialities.

Authors' contributions

Aviv Ladanie wrote the first draft with input from Lars G Hemkens and all authors made critical revisions to the manuscript. All authors read and approved the final version of the paper. Lars G Hemkens and Benjamin Kasenda obtained funding for this study. Aviv Ladanie and Lars G Hemkens are the guarantors.

Conflict of interest disclosures

All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf. All authors declare no financial relationships with any organization that might have an interest in the submitted work in the previous three years and no other relationships or activities that could appear to have influenced the submitted work.

Funding/ Support

This project was supported by the Swiss Cancer League (Grant No KLS-3587-02-2015). The Basel Institute for Clinical Epidemiology and Biostatistics is supported by Stiftung Institut für klinische Epidemiologie.

Role of the funder/sponsor

None of the funders/sponsors had a role in the design and conduct of the project and preparation, review, approval of the manuscript; and decision to submit the manuscript for publication.

Data sharing

No additional data available.

Copyright

Transparency declaration

The Corresponding Author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Ethical approval

Not required, this article does not contain any personal medical information about any identifiable living individuals.

References

1. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. N Engl J Med 2008;358(3):252-60.
2. Rising K, Bacchetti P, Bero L. Reporting bias in drug trials submitted to the Food and Drug Administration: review of publication and presentation. PLoS Med 2008;5(11):e217.
3. Lee K, Bacchetti P, Sim I. Publication of clinical trials supporting successful new drug applications: a literature analysis. PLoS Med 2008;5(9):e191-e.
4. Hart B, Lundh A, Bero L. Effect of reporting bias on meta-analyses of drug trials: reanalysis of

meta-analyses. BMJ (Clinical research ed). 2012;344:d7202-d.

5. Man-Son-Hing M, Wells G, Lau A. Quinine for nocturnal leg cramps. J Gen Intern Med 1998;13(9):600-6.

6. McDonagh MS, Peterson K, Balshem H, Helfand M. US Food and Drug Administration documents can provide unpublished evidence relevant to systematic reviews. J Clin Epidemiol 2013;66(10):1071-81.

7. MacLean CH, Morton SC, Ofman JJ, Roth EA, Shekelle PG. How useful are unpublished data from the Food and Drug Administration in meta-analysis? J Clin Epidemiol 2003;56(1):44-51.

8. Halfpenny NJA, Quigley JM, Thompson JC, Scott DA. Value and usability of unpublished data sources for systematic reviews and network meta-analyses. Evid Based Med 2016;21(6):208-13.

9. OpenTrialsFDA. OpenTrialsFDA: an interview with Erick Turner [Internet]. 10 May 2016. [cited 25 April 2018]. Available from: https://opentrials.net/2016/10/05/opentrialsfda-an-interview-with-erick-turner/.

10. Schroll JB, Bero L, Gotzsche PC. Searching for unpublished data for Cochrane reviews: cross sectional study. BMJ 2013;346(apr23 1):f2231-f.

11. Golder S, Loke YK, Wright K, Sterrantino C. Most systematic reviews of adverse effects did not include unpublished data. J Clin Epidemiol 2016;77:125-33.

12. O'Connor AB. The need for improved access to FDA reviews. JAMA 2009;302(2):191-.

13. Turner EH. How to access and process FDA drug approval packages for use in research. BMJ 2013;347(oct14 2):f5992-f.

14. Wolfe N, Gotzsche PC, Bero L. Strategies for obtaining unpublished drug trial data: a qualitative interview study. Syst Rev 2013;2:31.

15. Schroll JB, Abdel-Sattar M, Bero L. The Food and Drug Administration reports provided more data but were more difficult to use than the European Medicines Agency reports. J Clin Epidemiol 2015;68(1):102-7.

16. Goldacre B, Turner E. You can now search FDA approval documents easily at fda.opentrials.net.

BMJ 2017:j677-j.

17. Papathanasiou P, Brassart L, Blake P, Hart A, Whitbread L, Pembrey R, et al. Transparency in drug regulation: public assessment reports in Europe and Australia. Drug Discov Today 2016;21(11):1806-13.

18. The Cochrane Collaboration. What's new in the Cochrane Handbook? [Internet]. June 2017. [cited 25 April 2018]. Available from: http://training.cochrane.org/whats-new-cochrane-handbook.

19. Institute of Medicine (IOM). 2011. Data gathered as part of the FDA approval process. In. Washington, DC: The National Academies Press.

20. US Food and Drug Administration. Drugs@FDA database [Internet]. [cited 25 April 2018]. Available from: www.fda.gov/drugsatfda.

21. US Food and Drug Administration. Search for FDA guidance documents [Internet]. 15 May 2018. [cited 15 May 2018]. Available from: www.fda.gov/RegulatoryInformation/Guidances/default.htm.

22. US Food and Drug Administration. Manual of policies & procedures (CDER) [Internet]. 15 May 2018. [cited 21 May 2018]. Available from: www.fda.gov/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/ManualofPoliciesProcedures/default.htm.

23. US Food and Drug Administration. Good Review Practices (GRPs) [Internet]. 20 February 2018. [cited 25 April 2018]. Available from: www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/ucm118777.htm.

24. US Food and Drug Administration. Drugs@FDA frequently asked questions. What drug products are in Drugs@FDA? [Internet]. [cited 25 April 2018]. Available from: www.accessdata.fda.gov/scripts/cder/daf/index.cfm?event=faq.page#contains.

25. FDA medical review for triptorelin/TRELSTAR® (NDA20715).

26. FDA medical review for ustekinumab/STELARA® (BLA125261).

27. FDA medical review for pertuzumab/PERJETA® (BLA125409).

28. US Food and Drug Administration. Good review practice: statistical review template [Internet]. 30 July 2012 [cited 25 April 2018]. Available from: www.fda.gov/downloads/aboutfda/centersoffices/officeofmedicalproductsandtobacco/cder/manualofpoliciesprocedures/ucm313814.pdf.

29. US Food and Drug Administration. Freedom of information [Internet]. 28 March 2018. [cited 25 April 2018]. Available from: www.fda.gov/RegulatoryInformation/FOI/default.htm.

30. Tierney JF, Stewart LA, Ghersi D, Burdett S, Sydes MR. Practical methods for incorporating summary time-to-event data into meta-analysis. Trials 2007;8(1):16.

31. Parmar MKB, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. Stat Med 1998;17(24):2815-34.

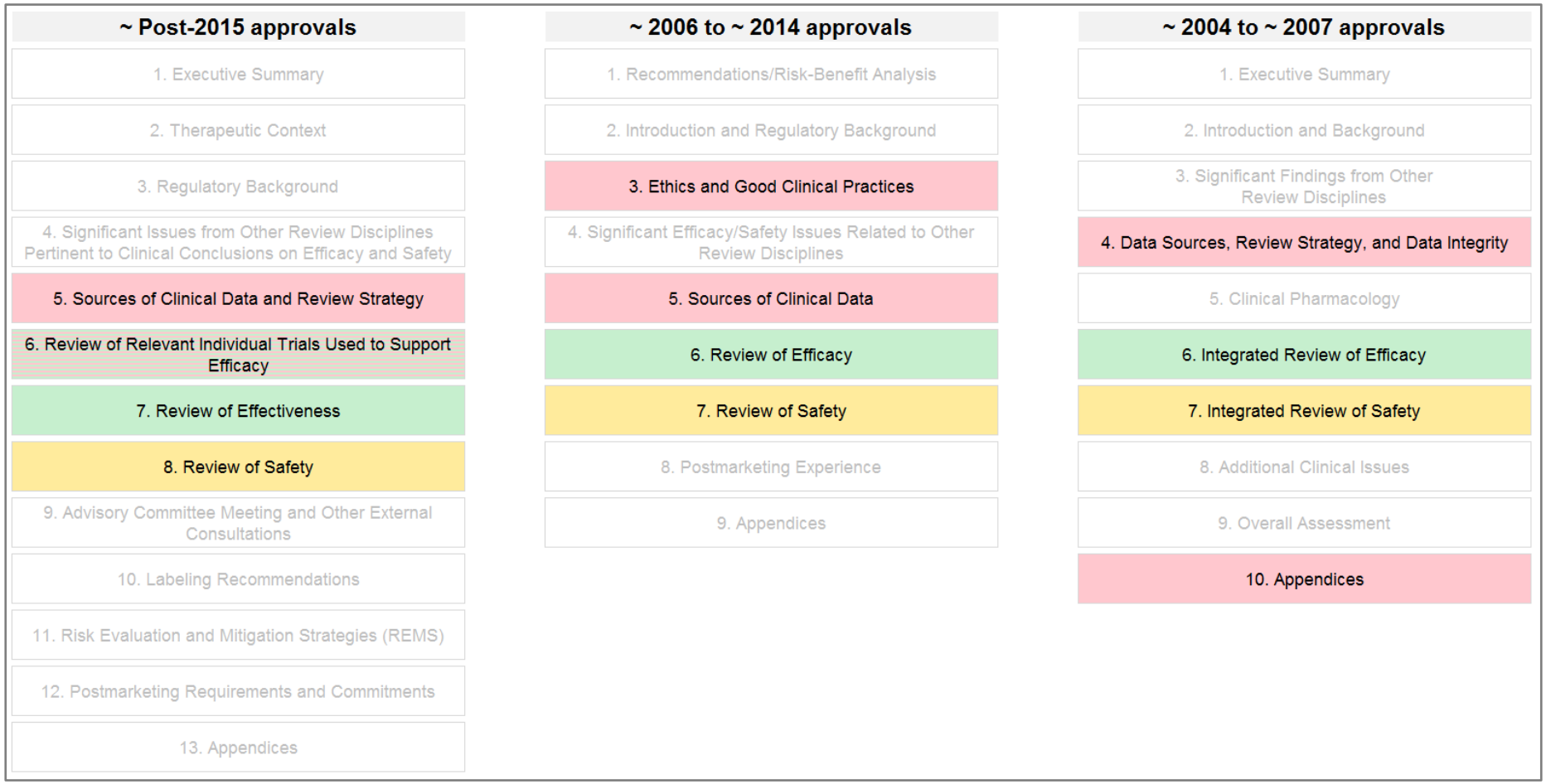
Box 1

Overview of documents that may be included in FDA approval packages:

1. Administrative documents
2. Approval letter
3. Botanical review
4. Chemistry, manufacturing, and controls review / Chemistry review
5. Clinical pharmacology and biopharmaceutics review / Clinical pharmacology review / Pharmacology review
6. Cross discipline team leader review
7. Drug label*
8. Environmental assessment review
9. Immunogenicity review
10. Medical review / Clinical review*
11. Multi-discipline review*
12. Medication guide / Patient package insert
13. Microbiology review
14. Name review
15. Office director memo
16. Officer/employee list
17. Other review
18. Risk assessment and risk mitigation review
19. Statistical review*
20. Summary review
21. Toxicology review

* Usually most relevant documents for evidence syntheses

FIGURE 1: OVERVIEW OF FDA MEDICAL REVIEW DOCUMENT STRUCTURES.



Sections with most relevant information for evidence syntheses are highlighted in colours. Pink sections: Provide an overview of the clinical trials submitted to the FDA and data on trial characteristics. Green sections: Describe trial results. Orange/Yellow sections: Report adverse events information. Greyed out sections: Typically less relevant for evidence syntheses of treatment effects of drugs and biologics. Note: Medical review documents of drugs approved prior to the year 2004 lack a consistent document structure and are not illustrated.

Box 2

Diversity of outcome analyses commonly seen in FDA medical reviews

- Intention-to-treat analyses versus per protocol or other analyses
- Different data cut-off-dates (for example interim analyses, final analyses, follow-up analyses)
- All study sites versus subsets of sites (for example geographic region)
- Local versus central outcome assessments
- Analyses conducted by the FDA versus analyses conducted by sponsors
- Analyses adjusted for covariates versus unadjusted analyses
- Pre-specified analyses versus post-hoc analyses

Box 3

Recommendations and some major points to consider

- As first step find out whether relevant FDA approval packages are available and how to access it
- When older documents are not searchable, consider using text recognition software
- The medical review is a key document and good starting point
- Try to locate the Table of Contents and get familiar with the overall document structure
- Identify pertinent trials and main information on trial characteristics and results in the medical review
- Note all trial names and identifiers, where possible
- Independent data extraction by multiple reviewers and noting the document pages where information was found may be helpful
- Consider further approval documents such as the statistical review, which may reveal important details
- Re-analyses by the FDA and comments of the different FDA reviewers may provide valuable insights for quality of evidence assessment.

- FDA reviewers have a unique view on the original research data that only a regulator can have
- Redacted information is common and requires special attention
- Compare data extracted from approval information with data from other sources
- Always assess the quality of evidence carefully, which sometimes requires to scrutinize the entire approval package to clarify unclear risks of bias

Supplemental material

(published as online data supplement in: BMJ 2018;362:k2815; https://doi.org/10.1136/bmj.k2815)

Part I: How to access FDA approval packages

In our main article, we describe how to use FDA approval documents for evidence syntheses in text form. Here we provide guidance on how to access and process FDA drug approval packages as an updated (and more detailed) version of the guide published in 2013 by EH Turner (Turner EH. How to access and process FDA drug approval packages for use in research. BMJ 2013;347:f5992. doi: https://doi.org/10.1136/bmj.f5992). An update is needed because the Drugs@FDA website (www.accessdata.fda.gov/scripts/cder/daf/) was restructured in 2016 (see rapid response by EH Turner, 19 March 2017 on www.bmj.com/content/347/bmj.f5992/rr-0).

The process is also outlined in a flowchart in Figure 1 in this document.

We illustrate the steps using the following working example:

Let us assume we are interested in the following PICOS question:

- Population: Men with metastatic hormone-refractory prostate cancer who have received prior chemotherapy
- Intervention: Cabazitaxel alone or in combination with prednisone
- Comparator: Any comparator without cabazitaxel
- Outcomes: Overall survival and health-related quality of life
- Study design: Randomized controlled trials

Step-by-step instruction

1. Where and how to find FDA approval documents

a) Determine if the compound of interest is approved by the US FDA's Center for Drug Evaluation and Research (CDER)

- Go to the Drugs@FDA start page (www.accessdata.fda.gov/scripts/cder/daf/)
- Enter the name of the compound of interest in the search bar (indicated by a red frame in Figure 2)
    - We recommend using the generic name instead of the brand name (e.g., Fluoxetine instead of PROZAC®), as the compound of interest is likely to be used in varied products with different names. For the next steps, we assume that the generic name was used for searching the database
- Click the search button located on the right side of the search bar
- If the compound of interest is regulated and approved by the FDA's CDER, you will be directed to an overview page presenting the search results
- If your search does not return any results, make sure you entered the:
    - Generic name (according to the United States Adopted Names, USAN, e.g., acetaminophen instead of paracetamol)
    - If your search remains unsuccessful, then the compound of interest is either a product not regulated by the FDA's CDER (e.g., a vaccine, such as GARDASIL®), or it has never been granted FDA approval in the US (e.g., domperidone).

b) Identify all products that may have approval packages

- If your search is successful, you may find yourself on a page which should look similar
    - to the one in Figure 5. In this case, there is only one product with the compound of interest existing, and you can directly jump to section c) Determine if there is at least one approval

package available

- to the one in Figure 3. In this case, multiple products are existing for the compound of interest

- Expand the panels pertaining to each product by clicking on them (note that that the search results page may have multiple pages to browse through). These newly visible entries (see Figure 4) are different applications submitted to the FDA for approval and are defined by a code (consisting of the type of application, i.e., New Drug Application, NDA, Abbreviated New Drug Application, ANDA, Biologics License Application, BLA), and a six-digit number assigned by the FDA
- Perform the following instructions for every listed product (otherwise you may falsely conclude that approval packages are unavailable). As a rule of thumb, you can ignore records labelled with:
  - ANDA: These are generic products approved under a different regulatory process for which no medical reviews are published or even existing, or
  - Tentative Approval: These are products that have not yet received final FDA approval, and therefore approval documents will not be available.

  Start with the first eligible product record by clicking on it (Figure 4). You will be directed to the product page (Figure 5).

c) Determine if there is at least one approval package available

- Expand the panel (indicated by a red frame in Figure 5) to see the history of regulatory actions
- In the green framed columns in Figure 6, search for “Review” entries (i.e., links to FDA approval packages) in the column “Letters, Reviews, Labels, Patient Package Insert”
  - Behind each “Review” link is an FDA approval package with information for a specific indication. If there are multiple reviews available, you should perform the following instructions for each “Review” link.

- Notice that approval packages are rarely available for approvals prior to 1997, for supplemental indications (as opposed to original indications, which are indications approved at the time when a drug receives its marketing authorization), or for drugs approved in the recent past (i.e., within the last couple of months).

> Working example
> - There is only one Review available (indicated by the red frame in Figure 6). This makes it highly likely that there is only one approved indication existing for the drug of interest.

- Click on the "Review" link to be directed to the approval package site (Figure 6)
- Download or open the PDF file by clicking on the "Medical Review(s)" Link (indicated by a red frame in Figure 7).

d) Do population and intervention characteristics match?

- Determine if the medical review is about the population and intervention that is of interest to you. Peruse the first couple of pages of the document to find a page giving an overview of the drug under review (e.g., brand name or therapeutic class), regulatory characteristics (e.g., date when the sponsor submitted the application to the FDA), as well as the intended target population of the drug
  - Some medical review documents may start with an addendum. In this case, there may be many more pages to scroll through until you find the information. As a rule of thumb, the information is available a few pages preceding the table of contents of the medical review.

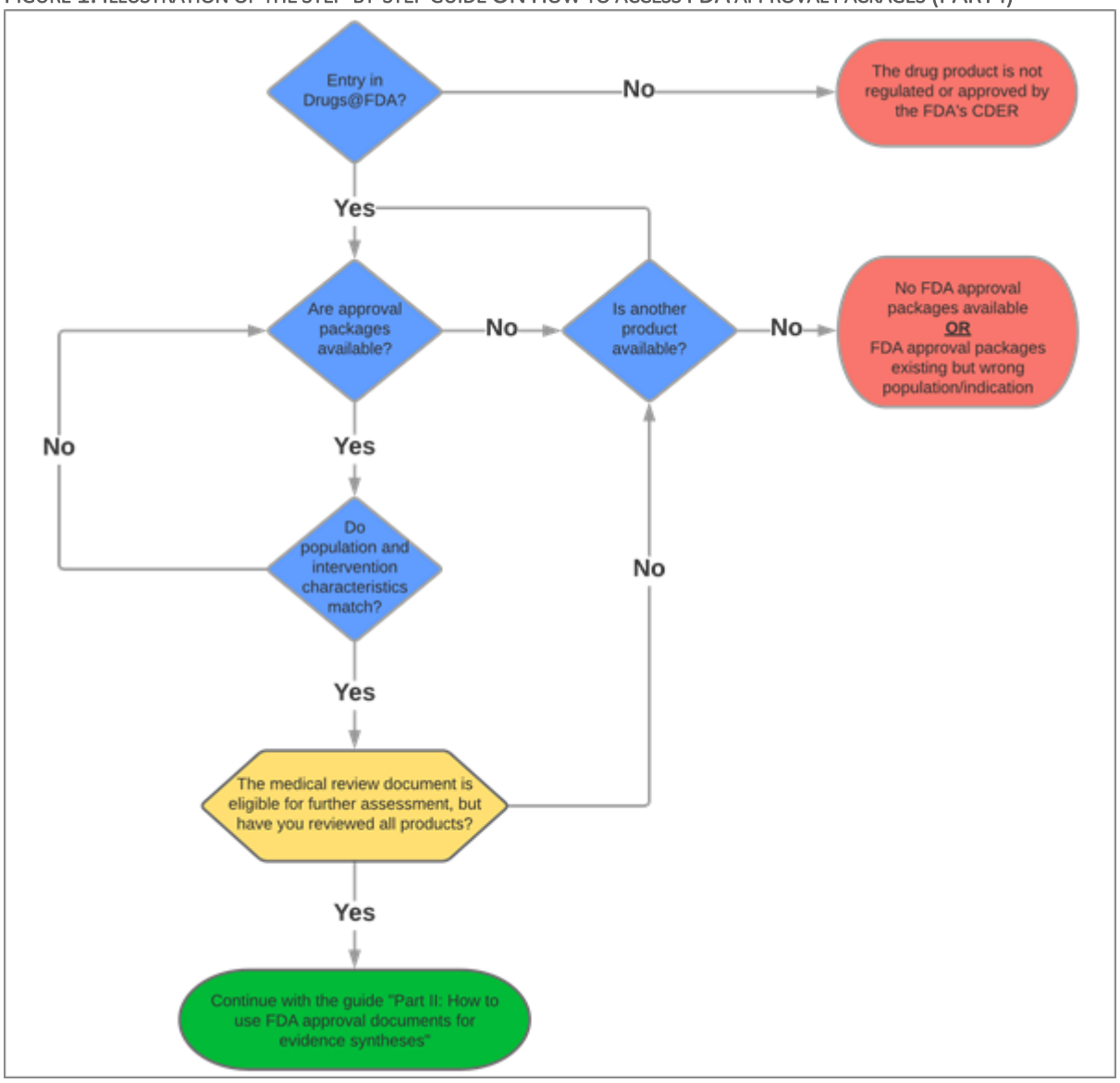
> Working example
> - This drug, regulatory and population information may be presented prominently over an entire page (such as in the case of cabazitaxel (see Figure 8) or only over a few paragraphs
> - The population and intervention for which cabazitaxel is indicated in this medical review matches our research question.

- The medical review is eligible for in depth assessment if the population and intervention match your research question.
- Examine if there are other indications or products existing that may also be relevant for your research question. For this purpose, repeat

- steps of section c) Determine if there is at least one approval package available to determine if there are other FDA approval packages available (for other but perhaps relevant indications), and, if there are none available, repeat subsequently
- steps of section b) Identify all products that may have approval packages to determine if there are other products (with the same compound of interest) approved that have FDA approval packages (Figure 1).

- If you have examined and obtained all potentially relevant medical reviews, continue with this process in "Part II: How to use FDA drug approval documents for evidence syntheses" of this document.

FIGURE 1: ILLUSTRATION OF THE STEP-BY-STEP GUIDE ON HOW TO ACCESS FDA APPROVAL PACKAGES (PART I)



Part II: How to use FDA drug approval documents for evidence syntheses

In our main article, we describe how to use FDA approval documents for evidence syntheses in text form. Here we illustrate the process step by step (which is also outlined in a flowchart in Figure 1) using the following working example:

Let us assume we are interested in the following PICOS question:

- Population: Men with metastatic hormone-refractory prostate cancer who have received prior chemotherapy
- Intervention: Cabazitaxel alone or in combination with prednisone
- Comparator: Any comparator without cabazitaxel
- Outcomes: Overall survival and health-related quality of life
- Study design: Randomized controlled trials

1. Where to find relevant clinical trials and their characteristics

- Find the table of contents (make sure that there is only one available; if there are two, perform all subsequent steps for each of them)
- Compare the structure of the table of contents (of the medical review document you have in front of you) with the three structures shown in Figure 1 of our main article, and find out which one is best matched

Working example

- The table of contents of cabazitaxel (see Figure 9) is located on pages 2 to 4
- The document structure is perfectly matched by the "~ 2006 to ~ 2014 approvals" structure illustrated in Figure 1 of our manual
- The chapters highlighted in blue ("3. Ethics and Good Clinical Practices", "5. Sources of Clinical Data", "6. Review of Efficacy", and "7. Review of Safety", see Figure 9) are the places where information about the clinical trials and their results are reported.

- Use the structure in Figure 1 of our main article to determine the best match to your medical document structure. The individual clinical trials are usually described in the chapter(s) framed in pink. Jump to the respective chapter in your medical review document
- Locate the overview of clinical studies, usually provided in tabular form

Working example

- According to the "~ 2006 to ~ 2014 approvals" structure illustrated in Figure 1 of our main article, the information is reported in the chapters framed in pink, i.e., either the chapter named "3 Ethics and Good Clinical Practices" or "5 Sources of Clinical Data"
  - Using the table of contents of our medical review document, we learn that these two chapters are on pages 17 and 21, respectively
  - After jumping to page 17 and perusing the chapter, we conclude that there is no tabular overview reported and continue our search in the second chapter (beginning on page 21)
- The "Table of Studies/Clinical Trials" is located on the page 22 of chapter 5 (see Figure 10).

- Assess eligibility of all studies presented in the overview and document your findings.

Working example

- There are five clinical trials listed in the table that are of potential interest to us (indicated by the red frames in Figure 10):
  - Only one study meets our inclusion criteria regarding the population (research question: men with metastatic hormone-refractory prostate cancer)
  - Four other studies (TED6188, TED6189, TED6190, BEX6702) may broadly be considered eligible based on the study population (solid tumours), but they appear to be early phase trials (dose escalation, bioavailability and pharmacokinetics studies) that are unlikely to meet our other eligibility criteria (be controlled trials and provide information on overall survival and health-related quality of life). However, in the absence of definite data, we will consider them eligible for the moment.

- Now that you know the clinical trials which are available in the medical review document, proceed with a) further determining study eligibility and b) extracting clinical trial information to describe the trial design. For this purpose, peruse the chapter(s) framed in pink in Figure 1 of our main article.

Working example

- After perusing chapter "3 Ethics and Good Clinical Practices", we conclude that no relevant clinical trial information is reported and continue with chapter "5 Sources of Clinical Data". After perusing this chapter, we conclude that only one clinical trial (EFC6193/TROPIC) was examined by the FDA for efficacy determination of cabazitaxel (as it is the only one that is described extensively). Information about the following characteristics can be found on the respective pages:

Characteristic	Page(s)	Comments
• Population	24 - 26	-
• Interventions	23	-
• Outcomes	23 - 24; 29 - 31	The list of endpoints tells us, that a) overall survival is the primary endpoint of the trial and b) that there is health-related quality of life measures are not available for this trial.
• Design	22 - 23	Statistical methods not reported in the medical review document but available from the statistical review on pages 6 - 9.

- Because the four clinical trials (TED6188, TED6189, TED6190, BEX6702) are not described in this chapter, we would expect that there is no sufficient information to critically understand and assess their findings. Nevertheless, now we know about their existence and a specific search in other approval documents and other sources of evidence (such as trial registries) may be helpful, and the trial identifiers may be very useful in this

regard.

2. Where to find trial results

- Use the structure in Figure 1 of our main article to determine the best match to your medical document structure.
  - Individual study results are usually presented and discussed in the chapter(s) framed in green (indicating where the trial conduct and efficacy results are discussed) and orange/yellow (indicating where safety results are discussed)
  - jump the respective chapter in your medical review document
- Determine if your outcome of interest is reported by perusing the sections, select the appropriate analysis (see Box 2 of our main article for an overview of the diversity of efficacy outcomes), and extract the information

Working example

- According to the "~ 2006 to ~ 2014 approvals" structure illustrated in Figure 1 of our main article, efficacy results are expected in chapter "6 Review of Efficacy" (indicated by one of the blue frames in Figure 9) on page 3. Information about the following efficacy characteristics can be found on the respective pages:

Characteristic	PDF page(s)	Comments
• Trial population	34 - 36	-
• Efficacy results		
o Overall survival	36 - 37	Only the applicant's analysis is reported* accompanied by a single sensitivity analysis.
o Health-related quality of life	-	Not available, as we have already noted earlier

* Including the number of patients analysed (per intervention group), number of patients with events (per intervention group), survival time (per intervention group), hazard ratio, p-value and Kaplan-Meier plot

- According to the "~ 2006 to ~ 2014 approvals" structure illustrated in Figure 1 of our main article, safety results are expected in chapter "7 Review of Safety" on page 43

- Remember to perform steps in section 1. Where to find relevant clinical trials and their characteristics and in section 2. Where to find trial results for each FDA approval document determined as eligible in section c) Determine if there is at least one approval package available. The total process is illustrated in Figure 1.

FIGURE 2: DRUGS@FDA – START PAGE (SCREENSHOT)
www.accessdata.fda.gov/scripts/cder/daf/

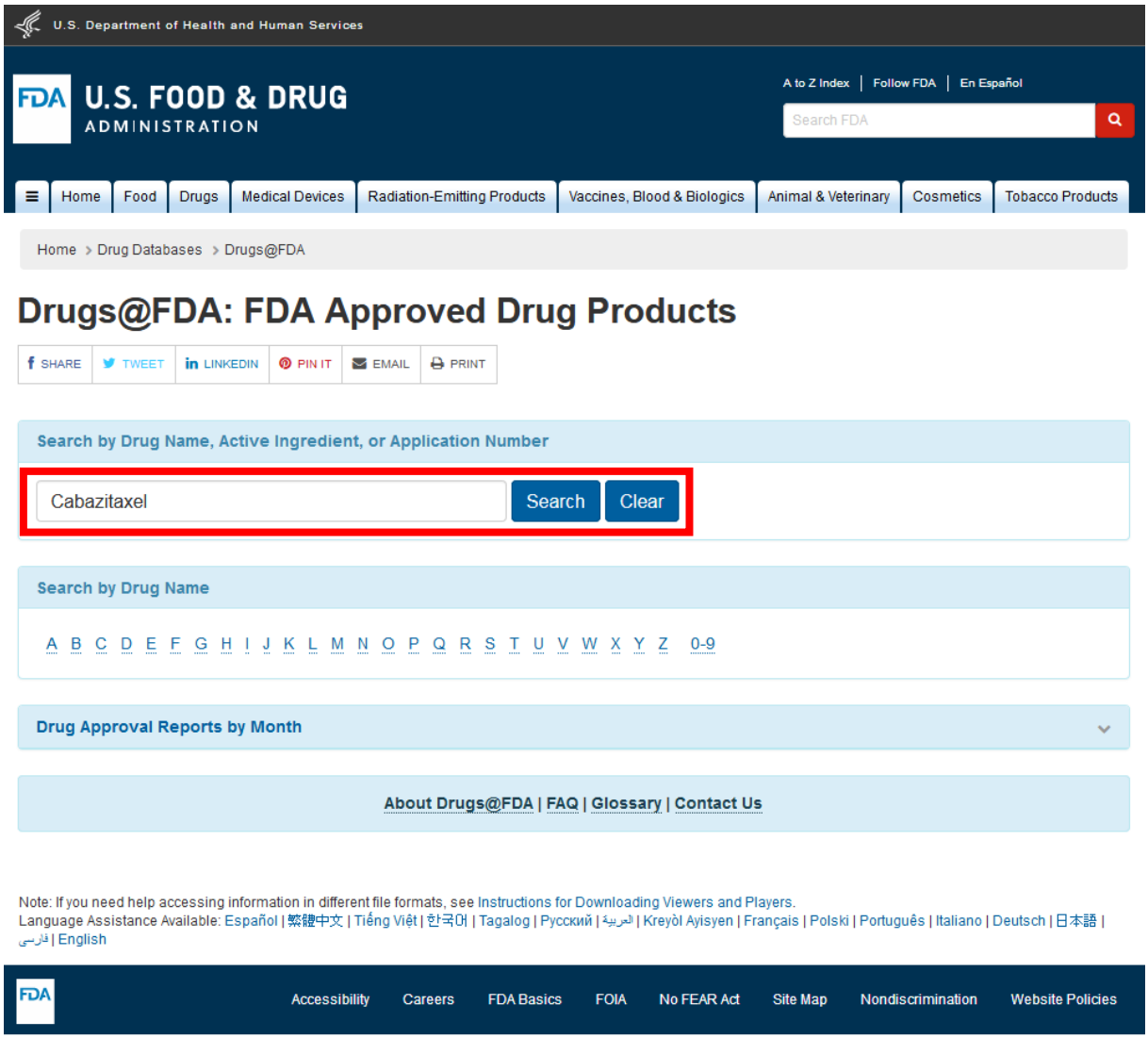


FIGURE 3: DRUGS@FDA – SEARCH RESULTS PAGE – PRODUCT OVERVIEW (1/2) (SCREENSHOT)

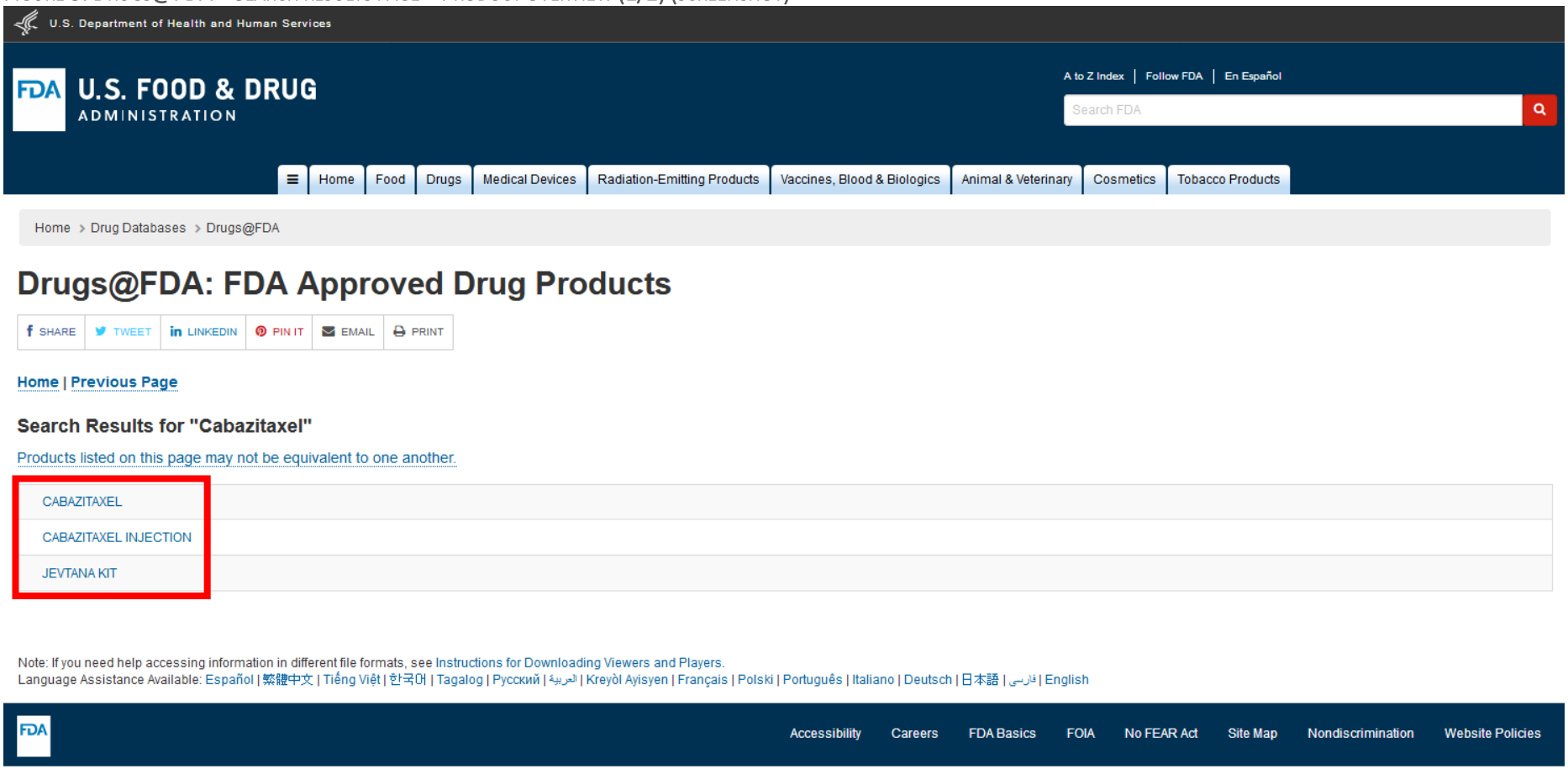


FIGURE 4: DRUGS@FDA – SEARCH RESULTS PAGE – PRODUCT OVERVIEW (2/2) (SCREENSHOT)

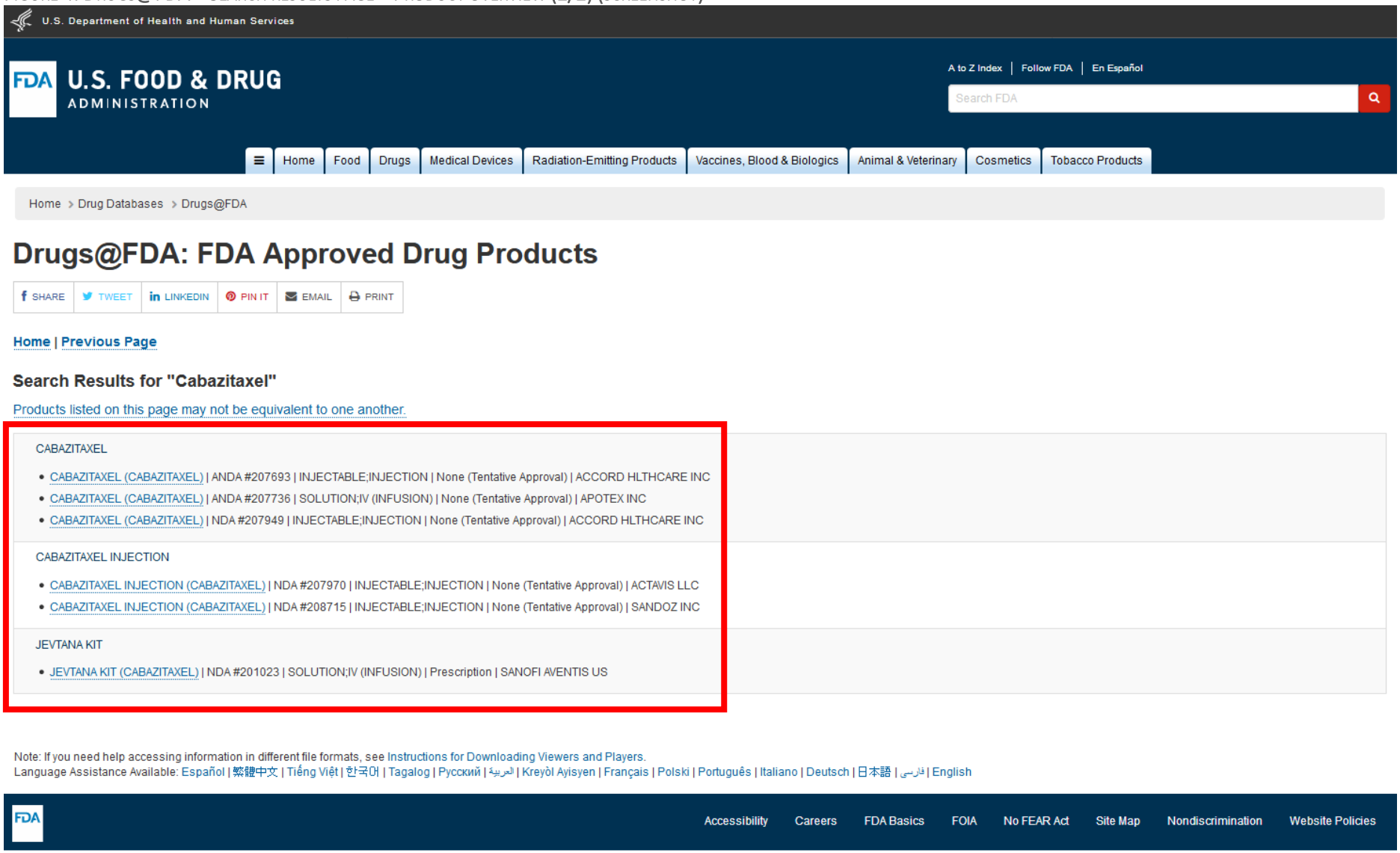


FIGURE 5: DRUGS@FDA – DRUG PRODUCT PAGE (1/2) (SCREENSHOT)

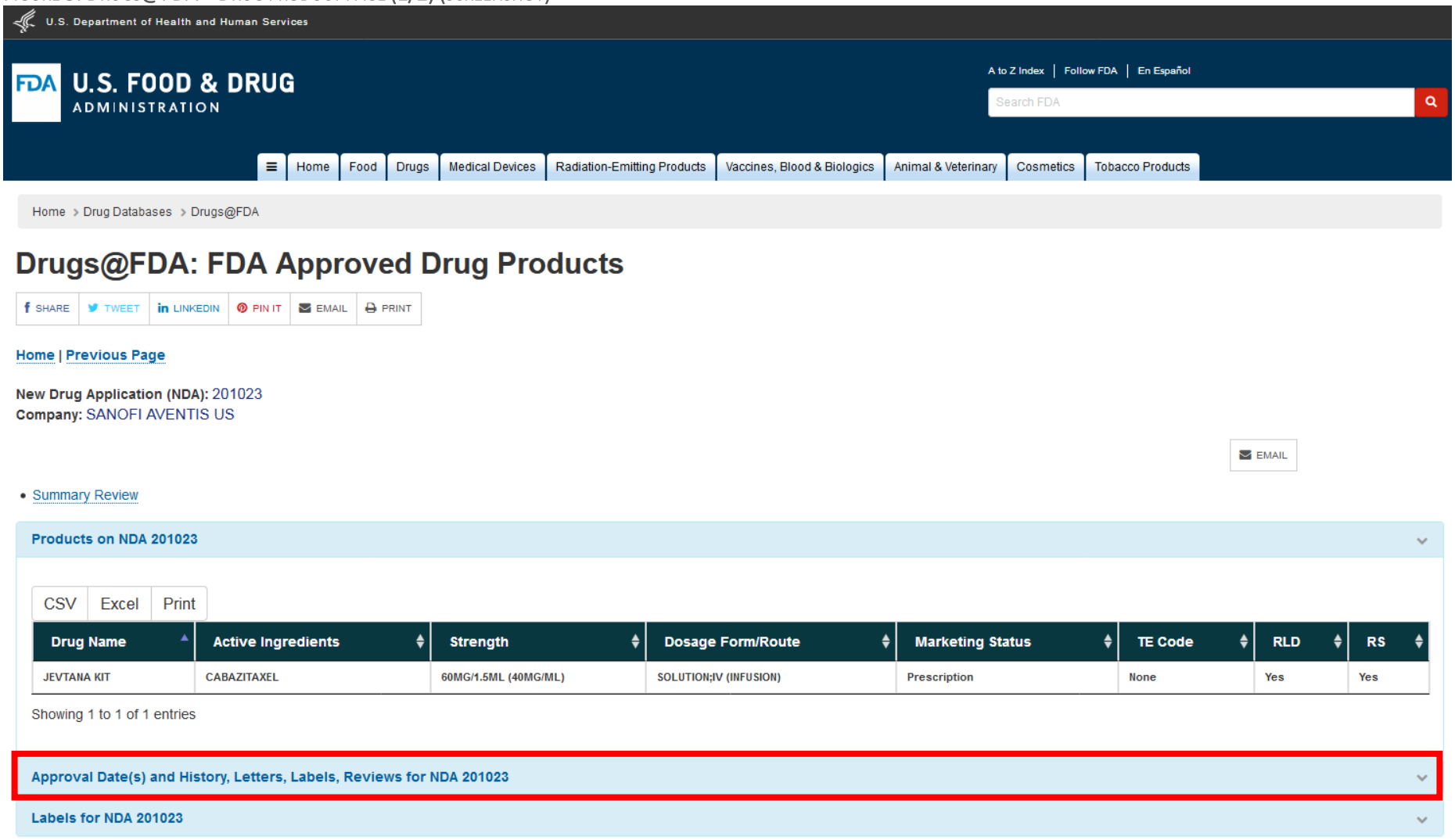
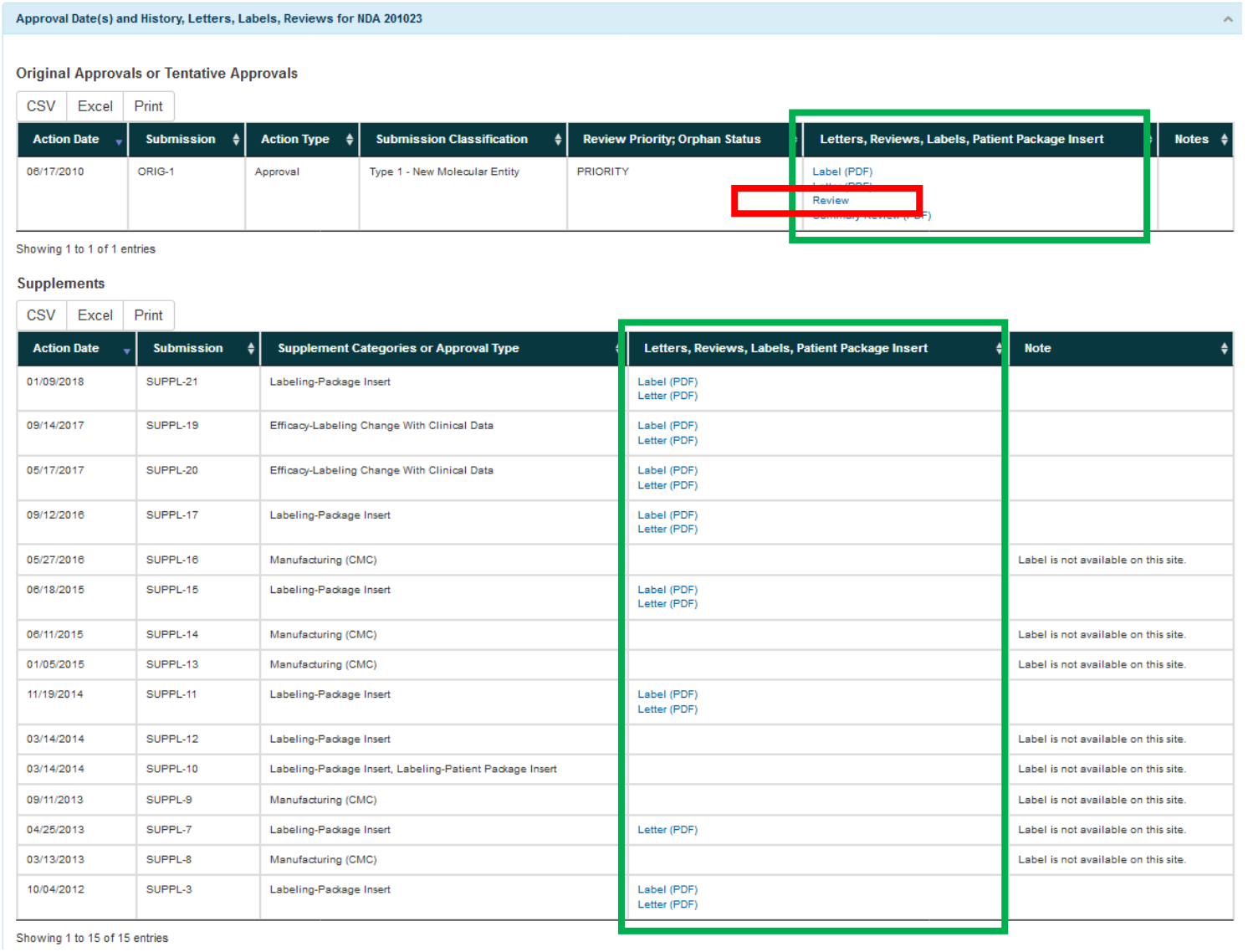


FIGURE 6: DRUGS@FDA – DRUG PRODUCT PAGE (2/2) (SCREENSHOT)



Approval Date(s) and History, Letters, Labels, Reviews for NDA 201023

Original Approvals or Tentative Approvals

CSV | Excel | Print

Action Date	Submission	Action Type	Submission Classification	Review Priority; Orphan Status	Letters, Reviews, Labels, Patient Package Insert	Notes
06/17/2010	ORIG-1	Approval	Type 1 - New Molecular Entity	PRIORITY	Label (PDF) Review	

Showing 1 to 1 of 1 entries

Supplements

CSV | Excel | Print

Action Date	Submission	Supplement Categories or Approval Type	Letters, Reviews, Labels, Patient Package Insert	Note
01/09/2018	SUPPL-21	Labeling-Package Insert	Label (PDF) Letter (PDF)	
09/14/2017	SUPPL-19	Efficacy-Labeling Change With Clinical Data	Label (PDF) Letter (PDF)	
05/17/2017	SUPPL-20	Efficacy-Labeling Change With Clinical Data	Label (PDF) Letter (PDF)	
09/12/2016	SUPPL-17	Labeling-Package Insert	Label (PDF) Letter (PDF)	
05/27/2016	SUPPL-16	Manufacturing (CMC)		Label is not available on this site.
06/18/2015	SUPPL-15	Labeling-Package Insert	Label (PDF) Letter (PDF)	
06/11/2015	SUPPL-14	Manufacturing (CMC)		Label is not available on this site.
01/05/2015	SUPPL-13	Manufacturing (CMC)		Label is not available on this site.
11/19/2014	SUPPL-11	Labeling-Package Insert	Label (PDF) Letter (PDF)	
03/14/2014	SUPPL-12	Labeling-Package Insert		Label is not available on this site.
03/14/2014	SUPPL-10	Labeling-Package Insert, Labeling-Patient Package Insert		Label is not available on this site.
09/11/2013	SUPPL-9	Manufacturing (CMC)		Label is not available on this site.
04/25/2013	SUPPL-7	Labeling-Package Insert	Letter (PDF)	Label is not available on this site.
03/13/2013	SUPPL-8	Manufacturing (CMC)		Label is not available on this site.
10/04/2012	SUPPL-3	Labeling-Package Insert	Label (PDF) Letter (PDF)	

Showing 1 to 15 of 15 entries

FIGURE 7: DRUGS@FDA – DRUG APPROVAL PACKAGE (SCREENSHOT)

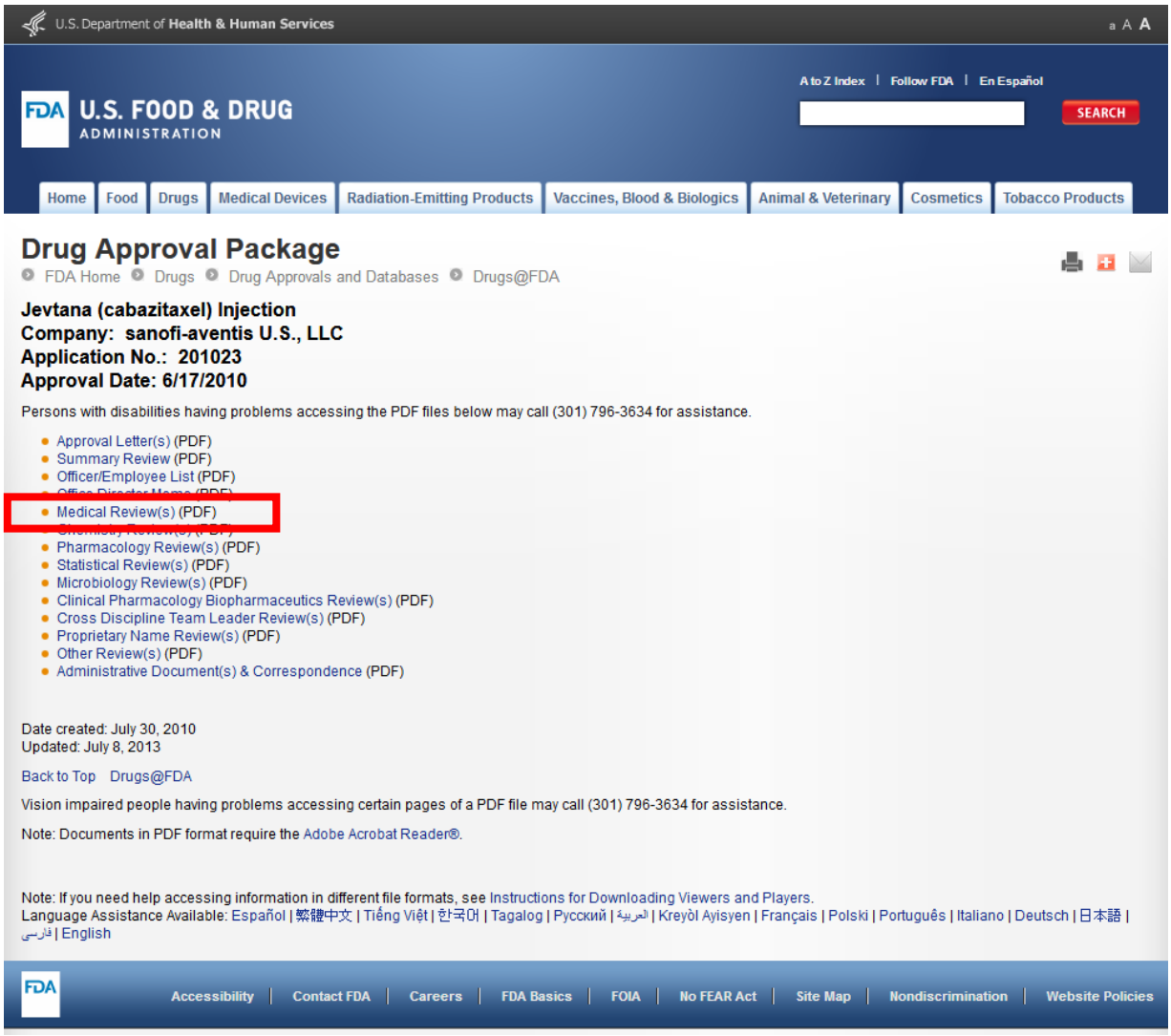


FIGURE 8: MEDICAL REVIEW – TITLE PAGE (SCREENSHOT)

CLINICAL REVIEW

Application Type	NDA
Application Number	201023
Priority or Standard	Priority
Submit Date	03/31/2010
Received Date	03/31/2010
PDUFA Goal Date	09/30/2010
Division / Office	DDOP / OODP
Reviewer Names	Amy McKee, M.D. Ian Waxman, M.D.
Review Completion Date	06/02/2010
Established Name	Cabazitaxel
Proposed Trade Name	Jevtana
Therapeutic Class	Taxane
Applicant	Sanofi-Aventis
Formulation	Intravenous
Dosing Regimen	25 mg/m^2 IV over 1 hour every 21 days
Indication	Hormone-refractory prostate cancer
Intended Population	Men with metastatic hormone-refractory prostate cancer who have previously received a docetaxel-containing regimen

FIGURE 9: MEDICAL REVIEW – TABLE OF CONTENTS (SCREENSHOT)

Table of Contents

2

3

FIGURE 10: MEDICAL REVIEW – TABLES OF STUDIES/CLINICAL TRIALS (SCREENSHOT)

5 Sources of Clinical Data

5.1 Tables of Studies/Clinical Trials

Table 1: Studies in NDA 201023

Study #	Population	Design	Dose[1] (mg/m^2)	# Any Cabazitaxel	# Cabazitaxel 25 mg/m^2 q 3 wks
TED6188	Solid Tumors	Dose Escalation	10-30 q 3 wks	21	6
TED6189	Solid Tumors	Dose Escalation	1.5-12 q 1 wk	31	0
		Oral Bioavailability	8.4 po c1d1 → 8.4 iv q wk	11	0
TED6190	Solid Tumors	Dose Escalation	10-25 q 3 wks	25	7
		Oral Biovailability	20 po c1d1 → 20 iv q 3 wks	11	0
BEX6702	Solid Tumors	PK Study	25 q 3 wks[2]	4	4
ARD6191	Metastatic Breast Cancer	Activity	10 q wk	13	0
			20-25 q 3 wks[3]	71	20
TCD6945	Metastatic Breast Cancer	Dose Escalation and Activity in Combination with Capecitabine	20-25 q 3 wks	33	6
EFC6193/ TROPIC	Metastatic Hormone-Refractory Prostate Cancer	Phase 3 Cabazitaxel + Prednisone vs. Mitoxantrone + Prednisone	25 q 3 wks	371	371
Total Exposed				**591**	**414**
ISS Total[4]				**558**	**408**
ISS					

4. The Comparative Effectiveness of Innovative Treatments for Cancer (CEIT-Cancer) project

Rationale and design of the database and the collection of evidence available at approval of novel drugs

Aviv Ladanie [1, 2], Benjamin Speich [1], Florian Naudet [3], Arnav Agarwal [4, 5], Tiago V Pereira [6], Francesco Sclafani [7], Juan Martin-Liberal [8, 9], Thomas Schmid [10], Hannah Ewald [1, 2, 11], John PA Ioannidis [12, 13, 14, 15, 16], Heiner C Bucher [1], Benjamin Kasenda [1, 17], Lars G Hemkens [1]

[1] Basel Institute for Clinical Epidemiology and Biostatistics, Department of Clinical Research, University Hospital and University of Basel, Basel, Switzerland
[2] Swiss Tropical and Public Health Institute (Swiss TPH), Basel, Switzerland
[3] Univ Rennes, CHU Rennes, Inserm, CIC 1414 [(Centre d'Investigation Clinique de Rennes)], F-35000 Rennes, France
[4] Department of Medicine, University of Toronto, Toronto, Ontario, Canada
[5] Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, Ontario, Canada
[6] Health Technology Assessment Unit, Institute of Education and Health Sciences, Oswaldo Cruz German Hospital, São Paulo, Brazil
[7] Department of Medicine, The Royal Marsden NHS Foundation Trust, Sutton, Surrey, UK
[8] Catalan Institute of Oncology (ICO) Hospitalet, Melanoma, Sarcoma and GU Tumors Unit, Barcelona, Spain
[9] Vall d'Hebron Institute of Oncology (VHIO), Early Drug Development Unit (UITM), Barcelona, Spain
[10] St. Clara Hospital, Basel, Switzerland
[11] University Medical Library, University of Basel, Basel, Switzerland
[12] Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, California, USA
[13] Department of Medicine, Stanford University School of Medicine, Stanford, California, USA
[14] Department of Health Research and Policy, Stanford University School of Medicine, Stanford, California, USA
[15] Department of Biomedical Data Science, Stanford University School of Medicine, Stanford, California USA
[16] Department of Statistics, Stanford University School of Humanities and Sciences, Stanford, California, USA
[17] Medical Oncology, University Hospital and University of Basel, Basel, Switzerland

Status: The final version of this manuscript is accepted for publication in: Trials.

Abstract

Background: The available evidence on the benefits and harms of novel drugs and therapeutic biologics at the time of approval is reported in publicly available documents provided by the United States Food and Drug Administration (FDA). We aimed to create a comprehensive database providing the relevant information required to systematically analyze and assess this early evidence in meta-epidemiological research.

Methods: We designed a modular and flexible database of systematically collected data. We identified all novel cancer drugs and therapeutic biologics approved by the FDA between 2000 and 2016, recorded regulatory characteristics, acquired the corresponding FDA approval documents, identified all clinical trials reported therein, and extracted trial design characteristics and treatment effects. Herein, we describe the rationale and design of the data collection process, particularly the organization of the data capturing, the identification and eligibility assessment of clinical trials, and the data extraction activities.

Discussion: We established a comprehensive database on the comparative effects of drugs and therapeutic biologics approved by the FDA over a time period of 17 years for the treatment of cancer (solid tumors and hematological malignancies). The database provides information on the clinical trial evidence available at the time of approval of novel cancer treatments. The modular nature and structure of the database and the data collection processes allow updates, expansions, and adaption for a continuous meta-epidemiological analysis of novel drugs.

The database allows to systematically evaluate benefits and harms of novel drugs and therapeutic biologics. It provides a useful basis for meta-epidemiological research on the comparative effects of innovative cancer treatments and continuous evaluations of regulatory developments.

Keywords: US Food and Drug Administration (FDA), Drug regulation, Marketing authorization, Approval package, Drugs and biologics, Clinical trials, Cancer, Evidence synthesis, Systematic review.

What is new

Key findings: We established a comprehensive database on novel cancer drugs and therapeutic biologics approved by the FDA between 2000 and 2016. The current database will be used to describe the clinical trial evidence generated in the pre-marketing period, but the database can and will be updated and expanded for future meta-epidemiological analyses.

What this adds to what is known: Publicly available drug approval documents offer highly valuable information that is very useful for evidence syntheses and research-on-research projects. The CEIT-Cancer database transparently describes and characterizes such information.

What is the implication, what should change now: This database allows systematically analyzing and assessing early evidence on benefits and harms of novel drug treatments and provides a solid basis for continuous meta-epidemiological analyses.

Introduction

Cancer drug development is characterized by an urgency to find novel treatments that improve patients' survival and quality of life. Timely access to such beneficial treatments is considered paramount for patients with cancer. Before granting approval and market access, health authorities such as the United States Food and Drug Administration (FDA) review the available evidence on benefits and harms from clinical trials and the claims made by the pharmaceutical companies and sponsors of the trials. The FDA examines the submitted clinical trial results, re-analyzes the trial's patient-level data and evaluates whether the trials were conducted and analyzed in accordance with the original study protocols (1, 2). For drugs and therapeutic biologics that receive approval, the FDA reviews are made publicly available in the drugs@FDA database as "approval packages" (3). These packages provide a wealth of information on the evidence on benefits and harms of innovative treatments at the time of approval.

With the introduction of new incentives and approval pathways, the FDA aimed to facilitate the development and approval process of drugs intended to treat serious or life-threatening conditions, including cancer (4). For example, some policies focus specifically on orphan drugs for rare diseases (4). Between 2000 and 2012, 46 of 47 oncology drugs approved by the FDA underwent expedited approval (5). In 2012, a further policy for so-called "breakthrough" therapies was introduced for drugs with highly promising pre-clinical evidence (5).

However, there is increasing discussion about the impact of these regulations because they may leave evidence gaps regarding efficacy and safety and increase uncertainty in clinical decision-making as expedited and orphan drug approvals are often based on smaller studies than used in traditional approvals (6). At the time of approval, there may be a dearth of evidence on hard clinical outcomes and subsequent follow-up evaluations suggest that such evidence may either never become available or may end up showing limited or no benefits (7-9). Oncology and hematology are probably the medical fields which are currently most affected by such developments.

Numerous meta-epidemiological studies aimed to better understand the evidence at the approval of novel cancer drugs and therapeutic biologics using data from the FDA and European Medicines Agency (EMA). We give an overview of these studies and the research in context in Table 1 (details of the underlying search strategy are provided in the Appendix). The first related investigation that we are aware of was published in 2009 (10), and the number of publications peaked in 2017 with 10 articles per year. Nonetheless, a major limitation is that many of these studies cover only certain types of cancer (e.g., solid tumors). Overall, there are four studies (10-13) which describe regulatory characteristics, clinical trials, assess endpoints and effect sizes used for approval on all cancer drugs, but none of them covers the most recently approved drugs (e.g., after 2013). This would not allow assessing newer policies such as the breakthrough program introduced in 2012. Thus, the current knowledge on approval evidence for cancer drugs is marked not only by a limited scope, but also a great diversity in methods and approaches, reducing the interpretability of the findings.

To address such limitations, we intended to establish a comprehensive database allowing for a

continuous analysis of such regulatory developments in meta-epidemiological research. The ongoing "Comparative Effectiveness of Innovative Treatments for Cancer" (CEIT-Cancer) project aims to transparently describe and characterize the clinical trial evidence of novel cancer drugs. Our goal is to capture the relevant information required to systematically analyze and assess early evidence on benefits and harms of novel cancer drug treatments.

As a first step, we collected the pre-marketing clinical trial evidence using FDA approval documents with a specific focus on cancer drugs, randomized controlled trials (RCTs) and single-arm trials (SATs), and treatment effects on overall survival (OS), progression-free survival (PFS), and objective tumor response (TR). However, the overall database structure is organized in a modular nature which allows for continuous updating of the list of drugs, the addition of new variables, expansion of the number of topics, health authorities, and outcomes, as well as linkage with other related datasets, e.g., from post-approval evidence including non-randomized real-world studies.

Herein, we describe the rationale and design of the data collection process for the pre-approval evidence, including the organization of the data capture, the identification of clinical trial information, the assessment of trials for eligibility, and the data extraction

Methods

Project organization and database structure

The data collection consisted of three steps: in step 1, we made an inventory of novel FDA approved drug products and acquired the corresponding FDA approval packages. In step 2, we made an inventory of RCTs and SATs reported in FDA approval documents, assessed their eligibility and extracted trial design characteristics. In step 3, we extracted treatment effects on OS, PFS, and TR.

Steps 2 and 3 started with a planning and organizing phase (operationalization of concepts, drafting of an instruction manual for standardized data selection and extraction, setting up the extraction

platform, pilot testing of the instruction manual and extraction platform, training of reviewers), followed by an execution phase (independent data extraction and verification) and ended with a closing phase (documentation of activities). Specific project activities are described in greater detail in the following sections.

The clinical trial data was managed in a single database. The database consists of four data tables (with information about the drug, indication, trial, study groups and treatment comparisons, as well as treatment effects) that are linked in one-to-many (1:n) relationships (Figure 1). The relational structure is indispensable because of the nature of the data (e.g., multiple indications approved for a single drug, multiple clinical trials supporting approval of a single indication, and multiple comparisons within a single multi-arm clinical trial). We used both Microsoft Access as a local data extraction and management platform and Ragic (www.ragic.com) as a cloud-based equivalent.

Step 1: inventory of FDA approved drugs and acquisition of approval packages

The aim of this step was to identify and characterize all drugs licensed by the FDA for the treatment of cancer diseases and to download as well as prepare FDA approval documents for subsequent activities. This step was performed by a single reviewer (AL).

Inventory of FDA approved drugs

In a first stage, we created a list of novel drugs and therapeutic biologics (referred to in this manuscript as “drugs”) that were granted their first FDA marketing authorization between 1 January 2000 and 31 December 2016 (technically speaking, we included so-called “new molecular entities” and “new therapeutic biologics” approved via either a “New Drug Application” or a “Biologics License Application”). The drug names were collected from the “Annual drug and biologic approval activity” reports for new molecular and biological entities (2000 to 2016) (14), as well as the “FDA reports on drug innovation” (2011 to 2016) (15). Information on therapeutic biologics approved before 2004 is not available in these documents and therefore, we reviewed the drug approval reports by month for the

period of January 2000 to December 2003 obtained from the Drugs@FDA database (3).

Selection of cancer indications

In a second stage, drugs were considered for inclusion in the CEIT-Cancer database if the original approval (i.e., the first ever approved use of a novel drug) was for the treatment of a solid tumor or hematological malignancy. Drugs without presumed cancer activity such as supportive care drugs (e.g., antiemetics, hematopoietic stem cell mobilizing agents) or imaging drugs (e.g., diagnostic radiopharmaceutical agents) were excluded. A medical oncologist (BK) was consulted in case of any doubts about eligibility.

Extraction of information on drug, indication and regulatory characteristics

In the third stage, we collated information on drug, indication and regulatory characteristics for each eligible drug and cancer indication ("drug-indication pair"; Table 2). The line of treatment was determined by a medical oncologist (BK). The remaining information was retrieved from various information sources as follows.

For drug-indication pair characteristics:

- "Annual drug and biologic approval activity" reports for new molecular and biological entities (2000 to 2016) (3), "FDA reports on drug innovation" (2011 to 2016) (15), and a peer-reviewed publication (16) for drug and regulatory characteristics, and
- the first-ever available FDA drug label from the drugs@FDA database (3) for information about the FDA approved indication(s).

For information on additional expedited programs and orphan status we perused:

- "FDA reports on accelerated approvals" to identify accelerated approved indications (17), i.e., indications approved based on preliminary evidence that does not meet regulatory standards for traditional (full) approval (4),
- "Breakthrough designation approval" reports (18) to identify indications that received a

breakthrough therapy designation in the pre-approval period, i.e., drugs that are expected to advance the treatment of certain diseases (4), and

- FDA database of orphan drug product designations to identify indications that received an orphan status (19), i.e., drugs intended for the treatment of rare diseases affecting less than 200'000 people in the United States (20).

All documents were downloaded or accessed on 02 November 2015 (for the 2000 to 2012 approvals) and on 02 March 2017 (for the 2013 to 2016 approvals). We relied on the information from the drugs@FDA database in the case of discrepant information between information sources (e.g., if there were different approval dates presented). We categorized the drug innovation class (first-in-class, advance-in-class, addition-to-class) according to the algorithm of Lanthier et al. (16). Accordingly, first-in-class drugs can be seen as "true" therapeutic innovation and define a new drug class. Advance-in-class drugs may offer important therapeutic advance (i.e., they were granted priority review by the FDA) over existing drugs in the same class. Drugs that do not fall in either of these two categories are categorized as addition-to-class.

Approval packages

The FDA's review of the preclinical and clinical information generated by a biopharmaceutical company during the course of drug development is summarized in FDA "approval packages" published in the drugs@FDA database. We used a similar approach to retrieve the approval documents as described recently (21), and we provided practical details on how we navigated the documents elsewhere (22). The following documents served as source documents throughout this project and were made suitable for text searching using Adobe Acrobat's Optical Character Recognition (OCR) function:

- Medical review (sometimes referred to as clinical review)
- Statistical review
- Drug label
- Cross-discipline team leader review

- Summary review
- Multi-discipline review.

Step 2: trial selection and characterization

The aim of this step was to identify eligible clinical trials in the medical review, assess their eligibility, and to characterize their design characteristics. These activities were performed by teams of two independent reviewers. Trials include randomized and non-randomized studies (the latter within the category of SATs), and for each trial the database explicitly indicates whether a randomized design was used.

Identification of trials, eligibility assessment, and data extraction

Each reviewer was provided with a set of indications to identify potentially eligible trials. Reviewers independently searched the medical review document for randomized trials as well as for trials that were indicated as pivotal for approval (i.e., the trial was described as “approval”, “registration”, “major”, “pivotal”, or similar), regardless of whether they were randomized or not. For each trial, the reviewers recorded variables presented in Table 3. In particular, they extracted the study identifier, name, or acronym, and determined if the following criteria were met (each criterion was assessed separately):

(1) the trial was explicitly described as pivotal to approval

(2) the patients were randomly assigned to treatment arms

(3) the patients matched broadly in their disease characteristics with the approved target population

(4) the patients were randomized to at least one control arm that did not contain the drug under review (regardless of dose or administration schedule)

(5) as per the judgment of the reviewer, a trial could still be relevant even if none of the abovementioned criteria were met, for example if the trial is extensively discussed or the only trial evaluated in the medical review (which is sometimes the case in accelerated approval settings,

where such trials are often not explicitly labeled as "pivotal" but extensively discussed in the documents).

After completion, the two independently generated datasets were compared, and disagreements resolved by consensus. The inter-rater reliability for trial identification (as assessed with the Kappa statistic (23)) was good (74%). Ultimately, trials that met either of the following set of criteria were deemed eligible:

- the trial was described as pivotal (criterion 1 alone is met; categorized as "explicitly pivotal")
- the trial was not described as pivotal, but was randomized (criterion 2), enrolled a population that matched the approved target population (criterion 3), and had a control arm that did not contain the intervention under review (criterion 4) (categorized as "likely pivotal RCT")
- the trial was not "explicitly pivotal" or a "likely pivotal RCT" but considered otherwise essential (criterion 5) for the approval decision (categorized as "other pivotal"). Such trials were typically single-arm studies in accelerated approval settings.

For each eligible trial, teams of two independent reviewers extracted information on variables presented in Table 4.

Step 3: treatment effect estimates on overall survival, progression-free survival, and tumor response

The aim of this step was to retrieve treatment effect estimates on OS, PFS, and TR for each treatment comparison. This information was collected only for RCTs. This activity was performed by teams of two independent reviewers.

Data extraction

We preferred trial analyses conducted by the FDA over sponsors' analyses, whenever both were available. Similarly, more recent data cutoff dates were preferred over older cutoff dates if there were multiple analysis results on the same endpoint available. We used the statistical review document (or

any other FDA approval documents) if the medical review document was not available, incomplete, or not legible.

For each treatment comparison, two reviewers independently searched the FDA review documents for treatment effect estimates on OS, PFS, and TR, and extracted information on variables presented in Table 5. For OS and PFS endpoints with incomplete or missing information (e.g., no confidence interval), we approximated treatment effect estimates following the methods described by Parmar et al. (24) and Tierney et al. (25).

At the end of the data collection activities in this step, the datasets of the two reviewers evaluating the same set of treatment comparisons were compared, and disagreements were resolved by consensus.

Discussion

We have successfully developed the CEIT-Cancer database, which transparently describes and characterizes information on the clinical trial evidence of novel cancer drugs at the time of their approval by the FDA.

Exploring characteristics of the evidence of novel cancer drugs at the time of their approval could greatly improve our understanding of the real-world clinical benefit and safety of such treatments. Importantly, it may also open new avenues of future research and regulation, leading to better-designed studies, reduced waste in research and more rigorous criteria for health authorities and health systems to consider incorporating new interventions to the current cancer armamentarium.

The CEIT-Cancer database is a comprehensive, manually curated platform that captures regulatory, drug, indication, and clinical trial data from FDA approvals of novel cancer drugs. This database differs from previous investigations in three important ways. First, the CEIT-Cancer database covers a time frame of 17 years, substantially larger compared to most previous studies. Second, it assesses all types of cancers, including both solid tumors and hematologic malignancies. Third, the database encompasses the most recent FDA drug approvals. In addition, this database can be expanded to other

medical fields and to be linked with other databases. It can be augmented with post-approval evidence and also allows expansion for being used for data extraction of approval documents from other health authorities, such as EMA (11, 26).

We have set-up the database and realized the project in a multidisciplinary team including experts in clinical trial methodology and conduct, clinical epidemiology, health technology assessment, biostatistics, clinical research, information management, public health, and medical oncology. The initial dataset covers a time period of 17 years. This allows us to investigate several regulatory developments over time and changes in the focus of drug development, such as the development of targeted agents and immunotherapy in contrast to classical cytotoxic chemotherapy. Following standardized and established data extraction procedures as in systematic reviews, we created a large evidence base on treatment effects and trial quality. This lays the foundation for our planned continuous meta-epidemiological analysis of novel drugs and therapeutic biologics within the CEIT-Cancer project. We currently develop the infrastructure to make the database available and aim to obtain structural funding and support to provide a sustainable solution. Through the collaborating participation of other investigators, we aim to establish a data sharing process to provide access to the database and foster further research.

Conclusion

Publicly available drug approval documents offer highly valuable information that is very useful for evidence syntheses and research-on-research projects. The CEIT-Cancer database transparently describes and characterizes this information on the clinical trial evidence of novel cancer drugs. It allows systematically analyzing and assessing early evidence on benefits and harms of novel drug treatments in meta-epidemiological research. Through the collaborating participation of other investigators, we aim to establish a data sharing process to provide access to the database and foster further research. The modular nature and structure of the database as well as the data collection processes permit

continuous updates and expansions. Overall, the database provides a solid basis for meta-epidemiological research of the evidence on novel treatments in cancer.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and material

The datasets generated and/or analyzed during the current study are not publicly available because data collection and processing is ongoing. They will be made available from the corresponding author on reasonable request as described in the article.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the Swiss Cancer League (KLS-3587-02-2015). The Meta-Research Innovation Center at Stanford is funded by a grant by the Laura and John Arnold Foundation. The Basel Institute of Clinical Epidemiology and Biostatistics is supported by Stiftung Institut für klinische Epidemiologie. None of the funders/sponsors had a role in the design and conduct of the project and preparation, review, approval of the manuscript; and decision to submit the manuscript for publication.

Authors' contributions

AL wrote the first draft with input by LGH, and all authors made critical revisions to the manuscript. All authors read and approved the final version of the paper. LGH and BK obtained funding for this study. AL and LGH are the guarantors.

Acknowledgements

Not applicable.

Tables and figures

TABLE 1: PUBLICATIONS WITH SIMILAR OR OVERLAPPING RESEARCH QUESTIONS

		Study characteristics				Characteristics described or analyzed			
(Reference, publication year)	Objective	Health authority	Time period	Approval type	Disease characteristics	Regulatory *	Trials	Endpoints	Effect sizes
Zeitoun (2018) (27)	To characterize post-marketing trials of cancer drugs	EMA, FDA	2005 - 2010	Original and supplemental	Solid tumors and hematologic malignancies	x	x	x	
Barnes and Amir (2017) (28)	To describe the efficacy, safety, tolerability, and price of new cancer drugs.	FDA	2005 - 2016	Not described/Unclear	Solid tumors only		x	x	x
Booth (2017) (29)	To evaluate the value of novel drugs using the ESMO Magnitude of Clinical Benefit Scale and ASCO Value Framework.	FDA	2015 - 2016	Not described/Unclear	Selected solid tumors [a]		x	x	
Brooks (2017) (30)	To understand the consequences of delaying approval of novel drugs until data on overall survival is available	FDA	1952 - 2016	Original and supplemental	Ten (10) most common solid tumors [b]			x	
Davis (2017) (11)	To determine the availability of data on overall survival and quality of life benefits of cancer drugs.	EMA	2009 - 2013	Original and supplemental	Solid tumors and hematologic malignancies	x	x	x	x
Grossmann (2017) (31)	To investigate the extent of European Medicines Agency (EMA)-approved cancer drugs that meet the threshold for “meaningful clinical benefit”, defined by the framework.	EMA	2011 - 2016	Original and supplemental	Solid tumors and hematologic malignancies		x	x	
Naci (2017) (32)	To characterize preapproval and confirmatory clinical trials of drugs granted accelerated approval.	FDA	2009 - 2013	Original and supplemental	Solid tumors and hematologic malignancies [c]		x	x	
Naci (2017) (9)	To systematically evaluate the timing and characteristics of clinical trials of drugs receiving accelerated approval.	FDA	2000 - 2013	Original only	Any disease or medical condition	x	x	x	
Pease (2017) (33)	To characterize controlled studies for drugs approved based on limited evidence.	FDA	2005 - 2012	Original only	Any disease or medical condition [d]	x	x	x	
Salas-Vega (2017) (34)	To evaluate the comparative therapeutic value of all new cancer medicines.	EMA, FDA	2003 - 2013	Original only	Solid tumors and hematologic			x	x

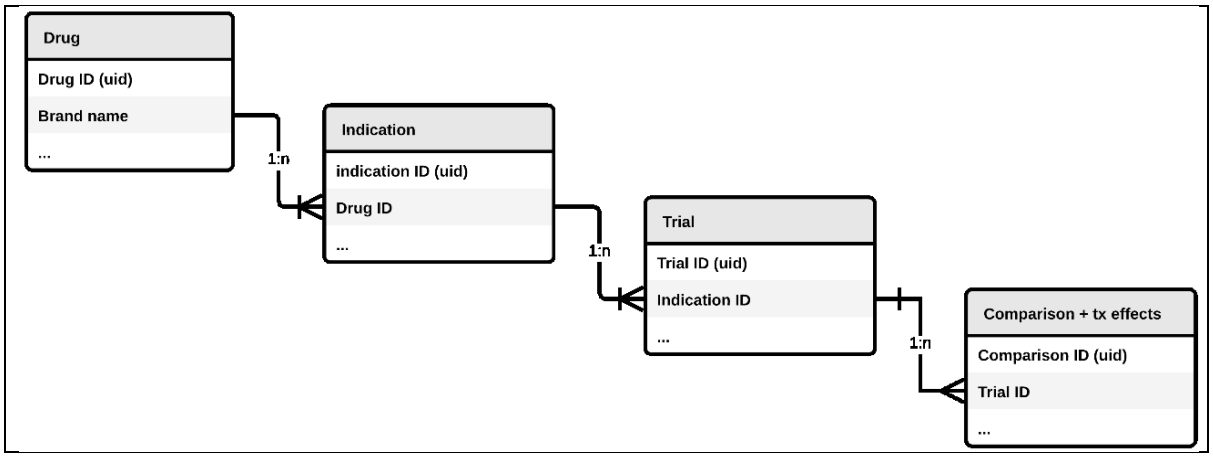
		Study characteristics				Characteristics described or analyzed			
(Reference, publication year)	Objective	Health authority	Time period	Approval type	Disease characteristics	Regulatory *	Trials	Endpoints	Effect sizes
					malignancies				
Smith (2017) (35)	To characterize the primary endpoints used to support United States FDA approvals for new drug or novel hematologic malignancies indications.	FDA	2002 - 2015	Original and supplemental	hematologic malignancies only	x	x	x	x
Tibau (2017) (36)	To derive the clinically meaningful benefit for FDA approved drugs using the "European Society for Medical Oncology - Magnitude of Clinical Benefit Scale".	FDA	2006 - 2016	Original and supplemental	Solid tumors only		x		
Grossmann (2016) (37)	To describe the knowledge about the clinical benefit of new cancer therapies at the time of approval.	EMA	2009 - 2016	Original and supplemental	Solid tumors and hematologic malignancies			x	x
Hoekman (2015) (38)	To describe the marketing authorization of oncology medicines granted based on the conditional marketing authorization pathway	EMA	2006 - 2013	Original only	Solid tumors only	x	x	x	
Kim (2015) (7)	To describe how often cancer drugs are approved based on a surrogate endpoint, whether subsequent studies for these drugs are reported, and whether the drugs improve overall survival.	FDA	2008 - 2012	Not described/Unclear	Solid tumors and hematologic malignancies	x		x	
Wang (2015) (39)	To characterize the types of comparators and endpoints used in efficacy trials for approvals of supplemental indications, compared with the data supporting these drugs' originally approved indications.	FDA	2005 - 2014	Supplemental only	Any disease or medical condition [e]	x	x	x	
Winstone (2015) (40)	To characterize the clinical trial evidence of orphan drugs.	EMA	2006 - 2014	Not described/Unclear	Solid tumors and hematologic malignancies [f]	x	x	x	
Downing (2014) (6)	To characterize pivotal efficacy trials for newly approved novel therapeutic agents.	FDA	2005 - 2012	Original only	Any disease or medical condition [e]	x	x	x	
Fojo (2014) (41)	To determine the availability of data on overall survival and quality of life benefits of cancer drugs.	FDA	2002 - 2014	Not described/Unclear	Solid tumors only				x
Hartmann (2013) (12)	To review the outcomes of marketing authorization applications for cancer drugs	EMA	2006 - 2011	Original only	Solid tumors and hematologic malignancies	x	x	x	x
Martell (2013) (42)	To describe approval trends and characteristics.	FDA	1949 - 2011	Original and	Solid tumors and	x	x	x	

		Study characteristics				Characteristics described or analyzed			
(Reference, publication year)	Objective	Health authority	Time period	Approval type	Disease characteristics	Regulatory *	Trials	Endpoints	Effect sizes
				supplemental	hematologic malignancies				
Thomas (2013) (43)	To describe pre- and post-approval availability of published comparative efficacy studies.	FDA	2000 - 2010	Original only	Any disease or medical condition [g]	x	x		
Goldberg (2011) (44)	To quantify the availability of comparative efficacy data for new molecular entities.	FDA	2000 - 2010	Original only	Any disease or medical condition	x	x		
Johnson (2011) (45)	To provide an overview of the regulatory history of accelerated approved oncology products.	FDA	1992 - 2010	Not described/Unclear	Solid tumors and hematologic malignancies [c]		x	x	
Kesselheim (2011) (46)	To define characteristics of orphan cancer drugs and their pivotal clinical trials and to compare these with non-orphan drugs.	FDA	2004 - 2010	Not described/Unclear	Solid tumors and hematologic malignancies [f]	x	x	x	
Ocana (2011) (47)	To determine if a difference in outcome between the experimental and control groups was detected that was equal to or greater than the value predefined in the protocol	FDA	2000 - 2010	Not described/Unclear	Solid tumors only			x	
Sridhara (2010) (13)	To conduct an overview of products that were reviewed by the FDA's Office of Hematology and Oncology Products for marketing approval and the regulatory actions taken during July 2005 to December 2007.	FDA	2005 - 2007	Original and supplemental	Solid tumors and hematologic malignancies	x	x	x	x
Tsimberidou (2009) (10)	To review the long-term safety and efficacy or cancer drugs approved without evidence from randomized trials.	FDA	1973 - 2006	Original only	Solid tumors and hematologic malignancies [h]	x	x	x	x

This list is based on a systematic search (appendix) but not intended to be exhaustive, as some relevant articles were brought to our attention by experts and could not be found with our limited search approach. *For example, approval pathways such as accelerated approval or orphan drug status. Other regulatory characteristics (such as approval times, approval probabilities, or availability of pediatric label information) are not considered here. Abbreviations: EMA, European Medicines Agency; FDA, Food and Drug Administration. [a] Limited to breast, lung, colorectal, or pancreatic cancers; [b] Limited to breast, colorectal, endometrial, gastric, liver, pancreatic, prostate and renal cancer as well as melanoma and non-small-cell lung cancer (NSCLC); [c] Limited to drugs approved under accelerated approval; [d] Includes any drug approved on the basis of a single pivotal trial, pivotal trials that used surrogate markers of disease as primary

endpoints or both; [e] Includes any drug; [f] Limited to orphan drugs only; [g] Limited to therapeutic biologics only; [h] Limited to drugs approved without evidence from randomized trials.

FIGURE 1: DATABASE STRUCTURE USED IN THE CEIT-CANCER PROJECT FOR DATA COLLECTION AND MANAGEMENT



Abbreviations: ID, identifier; tx, treatment effects; uid, unique identifier.

TABLE 2: Variables collected in step 1 for each cancer drug-indication pair

Variable	(Data type), data value or code	Description and further elaboration
Drug characteristics		
Brand name	(Character string)	As accepted by the FDA and used in the United States.
Generic name	(Character string)	According to United States Adopted Names.
Type of active compound	"NME"; "NBE"	NME (New Molecular Entity, i.e., a small molecule) or NBE (New Biologic Entity, i.e., a biologic product).
Date of marketing authorization	(Date)	Format: YYYY-MM-DD.
Innovation class	"First-in-class"; "Advance-in-class"; "Addition-to-class"	Drug innovation class, following the definitions and categories described by Lanthier et al. (16). New molecular or new biological entities are categorized as "First-in-class" if they define a new drug class, as "Advance-in-class" if they offer significant therapeutic advance (i.e., they were granted priority review by the FDA) over existing drugs in the same class, or "Addition-to-class" in any other case.
Indication characteristics		
FDA approved indication	(Character string)	Medical condition for which the drug of interest has been approved, according to the first-ever available FDA drug label.
Line of treatment	"1st"; "2nd"; "3rd"; "4th"	The clinical order the treatment is given
NDA/BLA number	(Integer)	FDA's Original New Drug Application (NDA) or Biologics License Application (BLA) number. A unique identifier assigned to each application for approval submitted to the FDA.
Site of disease	"Breast"; "Digestive"; "Gastrointestinal"; "Endocrine and Neuroendocrine"; "Genitourinary"; "Gynecologic"; Leukemia"; Lymphoma"; "Musculoskeletal"; "Neurologic"; "Other - Multicentric Castleman's	Cancers by body location/system (following the classification by the National Cancer Institute (www.cancer.gov/types/by-body-location).

Variable	(Data type), data value or code	Description and further elaboration
	Disease"; Other - Other"; "T-cell malignancies"; "Respiratory/Thoracic"; "Skin"	
Regulatory characteristics		
Priority review	"Standard"; "Priority"	Priority review is an expedited FDA review program for drugs that provide a significant improvement over existing therapies.
Accelerated approval	"Yes"; "No"	Expedited FDA approval pathway for drugs that a) treat serious conditions, b) provide a meaningful advantage over available therapies, and c) demonstrate effects on a surrogate endpoint that is reasonably likely to predict clinical endpoints. Accelerated approved drugs do not meet regulatory standards for traditional or full approval and are therefore required to provide evidence of clinical benefit in subsequent pivotal trials.
Breakthrough therapy designation	"Yes"; "No"	An expedited program at FDA introduced in 2012 for drugs that are a) intended to treat serious conditions, and b) provide preliminary clinical evidence of substantial improvement over existing therapies.
Orphan designation	"Yes"; "No"	A status assigned by the FDA to rare disease indications if less than 200'000 people in the United States are affected.

TABLE 3: Variables collected in step 2 for trials that were randomized or explicitly labeled as pivotal

Variable	(Data type), data value or code	Description and further elaboration
Trial characteristics (for any trial identified in step 2)		
Trial name reference	(Character string)	Reference trial name.
Trial name 1	(Character string)	Alternative trial name 1.
Trial name 2	(Character string)	Alternative trial name 2.
Pivotal	"Yes"; "No"	Trial eligibility criteria: the trial is described as "pivotal" (or similar).
Randomized	"Yes"; "No"; "Single-arm"	Trial eligibility criteria: patients are randomly assigned to treatment arms.
On-label	"Yes"; "No"; "Partially"; "Not reported"	Trial eligibility criteria: the drug of interest is tested in the approved indication.
Comparator	"Yes"; "No"; "Partially"; "Not reported"	Trial eligibility criteria: the control intervention does not contain the active component of the drug under review.
Relevance	"Yes"; "No"	Trial eligibility criteria: two reviewers consider that this trial was definitely used for approval, but none of the abovementioned eligibility criteria are met.
Eligible rationale	"explicitly pivotal"; "likely pivotal"; "other pivotal"; "not eligible"	The rationale for trial eligibility based on eligibility algorithm.

TABLE 4: Variables collected in step 2 for eligible trials only

Variable	(Data type), data value or code	Description and further elaboration
Trial characteristics (for any trial deemed eligible in step 2)		
Randomization	"Yes"; "No"	Random allocation of patients to trial arms
N arms	(Integer)	The number of trial arms.

Variable	(Data type), data value or code	Description and further elaboration
Trial characteristics (for any trial deemed eligible in step 2)		
Other trial characteristics	"Parallel"; "Cross-over"; "Uncontrolled/historic control"	Patients are randomized to a concurrent control ("Parallel") or to a sequence of treatments ("Cross-over").
Comparison characteristics		
Arm 1		
Type	"Experimental"; "Active"; "Placebo"; "No treatment"; "Dose-comparison"	In add-on trials, comparators were categorized as "active" whenever an intervention given on top of an active treatment (e.g., standard of care with or without placebo). Comparators were categorized as "No treatment" if "supportive therapy" or "usual care" was given which included a wide variety of treatments rather than a specific intervention.
Characteristics	(Character string)	All interventions in arm 1, including drug names, doses, and route of administration. Interventions used to avoid treatment-related complications were not recorded, such as pre-treatment with acetaminophen/diphenhydramine to reduce infusion reactions with intravenous infusion of therapeutic biologics, or antiemetics to reduce nausea and vomiting associated with certain chemotherapies.
Arm 2		
Type	"Active"; "Placebo"; "No treatment"; "Dose-comparison"; "Uncontrolled/historic control"	See "Arm 1" above.
Characteristics	(Character string)	See "Arm 1" above.

TABLE 5: Variables collected in step 3 for eligible randomized trials retrieved on comparison level

Variable	(Data type), data value or code	Description and further elaboration
Overall survival OR Progression-free survival		
Is the endpoint reported	"Yes"; "No"	.
Response criteria	(Character string)	**Progression-free survival only**: response criteria used to measure response to treatment
Number of patients in arm 1	(Integer)	Number of patients in arm 1 included in the endpoint analysis
Number of patients in arm 2	(Integer)	Number of patients in arm 2 included in the endpoint analysis
Number of events in arm 1	(Integer)	Number of patients with events in arm 1 included in the endpoint analysis
Number of events in arm 2	(Integer)	Number of patients with events in arm 2 included in the endpoint analysis
Hazard ratio: coverage probability	(Float)	Confidence level (1-alpha) in the endpoint analysis
Hazard ratio: point estimate	(Float)	Hazard ratio point estimate (selection rule: primary analysis according to the FDA, but longest follow-up)
Hazard ratio: lower confidence bound	(Float)	The lower bound of the confidence interval of the hazard ratio estimate
Hazard ratio: upper confidence bound	(Float)	The upper bound of the confidence interval of the hazard ratio estimate
Randomization ratio	"1:1"; "Not 1:1"	Randomization ratio, extracted for incomplete

Variable	(Data type), data value or code	Description and further elaboration
		endpoint effects to derive appropriate statistics (see (24, 25))
Regression p-value	(Float)	Regression p-value of the endpoint effect, extracted for incomplete endpoint effects to derive appropriate statistics (see (24, 25))
Test type	“1-sided”; "2-sided"; "Not reported"	One- or two-sided p-value, extracted for incomplete endpoint effects to derive appropriate statistics (see (24, 25))
Hazard rate in arm 1	(Float)	Hazard rate in arm 1, extracted for incomplete endpoint effects to derive appropriate statistics (see (24, 25))
Hazard rate in arm 2	(Float)	Hazard rate in arm 2, extracted for incomplete endpoint effects to derive appropriate statistics (see (24, 25))
Logrank observed minus expected events in arm 1	(Integer)	Logrank Observed minus Expected (O-E) events in arm 1 (endpoint analysis), extracted for incomplete endpoint effects to derive appropriate statistics (see (24, 25))
Logrank observed minus expected events in arm 2	(Integer)	Logrank Observed minus Expected (O-E) events in arm 2 (endpoint analysis), extracted for incomplete endpoint effects to derive appropriate statistics (see (24, 25))
Logrank variance	(Float)	Logrank variance (endpoint analysis), extracted for incomplete endpoint effects to derive appropriate statistics (see (24, 25))
Median survival time in arm 1: point estimate	(Float)	Median survival time (point estimate) in arm 1
Median survival time in arm 1: lower confidence bound	(Float)	The lower bound of the confidence interval of the median survival time in arm 1
Median survival time in arm 1: upper confidence bound	(Float)	The upper bound of the confidence interval of the median survival time in arm 1
Median survival time in arm 2	(Float)	Median survival time (point estimate) in arm 2
Median survival time in arm 2: lower confidence bound	(Float)	Lower bound of the confidence interval of the median survival time in arm 2
Median survival time in arm 2: upper confidence bound	(Float)	Upper bound of the confidence interval of the median survival time in arm 2
Time unit	"Days"; "Weeks"; "Months"; "Years"; "Not reported"	Time unit used to measure median survival improvement
Tumor response		
Is the endpoint reported	"Yes"; "No"	./.
Primary endpoint	"Yes"; "No"	Is the tumor response endpoint described as the primary endpoint of the trial
Type of hypothesis tested	"Superiority"; "Not 1° endpoint"; "Non-inferiority"	Is the trial designed to demonstrate the superiority of the test drug over control in tumor response
Response criteria	(Character)	Set of response criteria used to measure tumor response
Number of patients in arm 1	(Integer)	Number of patients in arm 1 included in the tumor response endpoint analysis
Number of patients in arm 2	(Integer)	Number of patients in arm 2 included in the tumor response endpoint analysis
Number of events in arm 1	(Integer)	Number of patients with events in arm 1 included in the tumor response endpoint analysis

Variable	(Data type), data value or code	Description and further elaboration
Number of events in arm 2	(Integer)	Number of patients with events in arm 2 included in the tumor response endpoint analysis

References

1. Turner E. Correction/clarification about FDA review documents. PLoS Med 2005;2(12):e422; author reply.
2. US Food and Drug Administration. The drug development process - Step 4: FDA drug review [Internet]. 04 January 2018. [cited 04 June 2018]. Available from: www.fda.gov/ForPatients/Approvals/Drugs/ucm405570.htm.
3. US Food and Drug Administration. Drugs@FDA: FDA approved drug products [Internet]. [cited 04 June 2018]. Available from: www.fda.gov/drugsatfda.
4. US Food and Drug Administration. Guidance for industry: expedited programs for serious conditions – drugs and biologics [Internet]. May 2014. Available from: www.fda.gov/downloads/Drugs/Guidances/UCM358301.pdf.
5. Darrow JJ, Kesselheim AS. Drug development and FDA approval, 1938-2013. N Engl J Med 2014;370(26):e39.
6. Downing NS, Aminawung JA, Shah ND, Krumholz HM, Ross JS. Clinical trial evidence supporting FDA approval of novel therapeutic agents, 2005-2012. JAMA 2014;311(4):368-77.
7. Kim C, Prasad V. Cancer drugs approved on the basis of a surrogate end point and subsequent overall survival: an analysis of 5 years of US Food and Drug Administration approvals. JAMA Intern Med 2015;175(12):1992-4.
8. Kim C, Prasad V. Strength of validation for surrogate end points used in the US Food and Drug Administration's approval of oncology drugs. Mayo Clin Proc 2016;91(6):713-725.
9. Naci H, Wouters OJ, Gupta R, Ioannidis JPA. Timing and characteristics of cumulative evidence available on novel therapeutic agents receiving Food and Drug Administration accelerated

approval. Milbank Q 2017;95(2):261-90.

10. Tsimberidou AM, Braiteh F, Stewart DJ, Kurzrock R. Ultimate fate of oncology drugs approved by the us food and drug administration without a randomized Trial. J Clin Oncol 2009;27(36):6243-50.

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

12. Hartmann M, Mayer-Nicolai C, Pfaff O. Approval probabilities and regulatory review patterns for anticancer drugs in the European Union. Crit Rev Oncol Hematol 2013;87(2):112-21.

13. Sridhara R, Johnson JR, Justice R, Keegan P, Chakravarty A, Pazdur R. Review of oncology and hematology drug product approvals at the US Food and Drug Administration between July 2005 and December 2007. J Natl Cancer Inst 2010;102(4):230-43.

14. US Food and Drug Administration. New Molecular Entity (NME) drug and new biologic approvals [Internet]. 01 February 2018. [cited 04 June 2018]. Available from: www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/DrugandBiologicApprovalReports/NDAandBLAApprovalReports/ucm373420.htm.

15. US Food and Drug Administration. New Drugs at FDA: CDER's new molecular entities and new therapeutic biological products [Internet]. 02 February 2018. [cited 04 June 2018]. Available from: www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugInnovation/default.htm.

16. Lanthier M, Miller KL, Nardinelli C, Woodcock J. An improved approach to measuring drug innovation finds steady rates of first-in-class pharmaceuticals, 1987-2011. Health Aff (Millwood) 2013;32(8):1433-9.

17. US Food and Drug Administration. Accelerated approvals [Internet]. 30 January 2015. [cited 04 June 2018]. Available from: www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/DrugandBiologicApprovalReports/NDAandBLAApprovalReports/ucm373430.htm.

18. US Food and Drug Administration. Breakthrough therapy approvals [Internet]. 16 May 2017. [cited 04 June 2018]. Available from: www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/DrugandBiologicApprovalReports/NDAandBLAApprovalReports/ucm373418.htm.

19. US Food and Drug Administration. Search orphan drug designations and approvals [Internet]. [cited 04 June 2018]. Available from: www.accessdata.fda.gov/scripts/opdlisting/oopd/.

20. Institute of Medicine. 2010. Rare diseases and orphan products: accelerating research and development. Washington, DC: The National Academies Press. https://doi.org/10.17226/12953.

21. Turner EH. How to access and process FDA drug approval packages for use in research. BMJ 2013;347:f5992.

22. Ladanie A, Ewald H, Kasenda B, Hemkens LG. How to use FDA drug approval documents for evidence syntheses. BMJ 2018;362:k2815.

23. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-82.

24. Parmar MKB, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. Stat Med 1998;17(24):2815-34.

25. Tierney JF, Stewart LA, Ghersi D, Burdett S, Sydes MR. Practical methods for incorporating summary time-to-event data into meta-analysis. Trials 2007;8:16.

26. Djulbegovic B, Glasziou P, Klocksieben FA, Reljic T, VanDenBergh M, Mhaskar R, et al. Larger effect sizes in nonrandomized studies are associated with higher rates of EMA licensing approval. J Clin Epidemiol 2018;98:24-32.

27. Zeitoun JD, Baron G, Vivot A, Atal I, Downing NS, Ross JS, et al. Post-marketing research and its outcome for novel anticancer agents approved by both the FDA and EMA between 2005 and 2010: A cross-sectional study. Int J Cancer 2018;142(2):414-23.

28. Barnes TA, Amir E, Templeton AJ, Gomez-Garcia S, Navarro B, Seruga B, et al. Efficacy, safety, tolerability and price of newly approved drugs in solid tumors. Cancer Treat Rev 2017;56:1-7.

29. Booth CM, Del Paggio JC. Approvals in 2016: questioning the clinical benefit of anticancer

therapies. Nat Rev Clin Oncol 2017;14(3):135-6.

30. Brooks N, Campone M, Paddock S, Shortenhaus S, Grainger D, Zummo J, et al. Approving cancer treatments based on endpoints other than overall survival: an analysis of historical data using the PACE Continuous Innovation Indicators (CII). Drugs Context 2017;6:212507.

31. Grossmann N, Del Paggio JC, Wolf S, Sullivan R, Booth CM, Rosian K, et al. Five years of EMA-approved systemic cancer therapies for solid tumours-a comparison of two thresholds for meaningful clinical benefit. Eur J Cancer 2017;82:66-71.

32. Naci H, Smalley KR, Kesselheim AS. Characteristics of preapproval and postapproval studies for drugs granted accelerated approval by the US Food and Drug Administration. JAMA 2017;318(7):626-36.

33. Pease AM, Krumholz HM, Downing NS, Aminawung JA, Shah ND, Ross JS. Postapproval studies of drugs initially approved by the FDA on the basis of limited evidence: systematic review. BMJ 2017;357:j1680.

34. Salas-Vega S, Iliopoulos O, Mossialos E. Assessment of overall survival, quality of life, and safety benefits associated with new cancer medicines. JAMA Oncol 2017;3(3):382-90.

35. Smith BD, DeZern AE, Bastian AW, Durie BGM. Meaningful endpoints for therapies approved for hematologic malignancies. Cancer 2017;123(10):1689-94.

36. Tibau A, Molto C, Ocana A, Templeton AJ, Del Carpio LP, Del Paggio JC, et al. Magnitude of clinical benefit of cancer drugs approved by the US Food and Drug Administration. J Natl Cancer Inst 2017;110(5):486-492

37. Grossmann N, Wild C. Between January 2009 and April 2016, 134 novel anticancer therapies were approved: what is the level of knowledge concerning the clinical benefit at the time of approval? ESMO Open 2016;1(6):e000125.

38. Hoekman J, Boon WP, Bouvy JC, Ebbers HC, de Jong JP, De Bruin ML. Use of the conditional marketing authorization pathway for oncology medicines in Europe. Clin Pharmacol Ther 2015;98(5):534-41.

39. Wang B, Kesselheim AS. Characteristics of efficacy evidence supporting approval of supplemental indications for prescription drugs in United States, 2005-14: systematic review. BMJ 2015;351:h4679.

40. Winstone J, Chadda S, Ralston S, Sajosi P. Review and comparison of clinical evidence submitted to support European Medicines Agency market authorization of orphan-designated oncological treatments. Orphanet J Rare Dis 2015;10:139.

41. Fojo T, Mailankody S, Lo A. Unintended consequences of expensive cancer therapeutics-the pursuit of marginal indications and a me-too mentality that stifles innovation and creativity: the John Conley Lecture. JAMA Otolaryngol Head Neck Surg 2014;140(12):1225-36.

42. Martell RE, Sermer D, Getz K, Kaitin KI. Oncology drug development and approval of systemic anticancer therapy by the U.S. Food and Drug Administration. Oncologist 2013;18(1):104-11.

43. Thomas RH, Freeman MK, Hughes PJ. Preapproval and postapproval availability of published comparative efficacy research on biological agents. Am J Health Syst Pharm 2013;70(14):1250-5.

44. Goldberg NH, Schneeweiss S, Kowal MK, Gagne JJ. Availability of comparative efficacy data at the time of drug approval in the United States. JAMA 2011;305(17):1786-9.

45. Johnson JR, Ning YM, Farrell A, Justice R, Keegan P, Pazdur R. Accelerated approval of oncology products: the food and drug administration experience. J Natl Cancer Inst 2011;103(8):636-44.

46. Kesselheim AS, Myers JA, Avorn J. Characteristics of clinical trials to support approval of orphan vs nonorphan drugs for cancer. JAMA 2011;305(22):2320-6.

47. Ocana A, Tannock IF. When are "positive" clinical trials in oncology truly positive? J Natl Cancer Inst 2011;103(1):16-20.

Appendix

Literature search for research in context

We searched PubMed with two strategies because we realized that the first search was not able to capture several pertinent articles. But even with the second strategy we have not identified some articles that were brought to our attention by experts. This is a pragmatic search approach as we did not hand-search articles, used additional databases or citation-search techniques because this would be beyond the scope of this supplementary information. We also used cancer-specific terms in the search strategy and focused on FDA and EMA. Overall, we aimed to capture most related articles and give an overview, but this can't replace a complete systematic review with the intention to provide an exhaustive list of similar articles.

Titles and abstracts of references identified with the main search strategy were screened independently by two reviewers (LGH, BK). The titles of references identified with the second search were screened by one reviewer (AL). Full texts of all references deemed potentially eligible by any of the three reviewers in any search were assessed for eligibility by teams of two reviewers (AL, BK, or LGH).

We included articles describing a systematic analysis of approval evidence (general or cancer-specific) which reported specific analyses for oncological drugs and included a description of characteristics of (1) regulatory circumstances; (2) clinical trials; (3) endpoints; (4) or treatment effects.

First Search

Search date: 27 December 2017: 480 hits. Search terms:

("Food and Drug"[tiab] OR United States Food and Drug Administration[mh] OR drugs@fda[tiab] OR "European medicines agency"[tiab] OR "Marketing Authorizations"[ti] OR "Marketing Authorisations"[ti] OR "Marketing Authorization"[ti] OR "Marketing Authorisation"[ti])

AND (oncology[tw] OR oncological[ti] OR "neoplasms"[MeSH Terms] OR neoplas*[ti] OR cancer*[ti] OR malignan*[ti] OR tumor[ti] OR tumour[ti] OR cancer[sb] OR anticancer[ti])

AND (APPROV*[tiab] OR PREAPPROVAL[tiab] OR Drug Approval[mh] OR PIVOT*[tiab])

AND (compar*[ti] OR Cross-Sectional[ti] OR Cross-Sectional Studies[mh] OR (Between January 19*[ti] OR Between February 19*[ti] OR Between March 19*[ti] OR Between April 19*[ti] OR Between May 19*[ti] OR Between June 19*[ti] OR Between July 19*[ti] OR Between August 19*[ti] OR Between September 19*[ti] OR Between October 19*[ti] OR Between November 19*[ti] OR Between December 19*[ti] OR Between January 20*[ti] OR Between February 20*[ti] OR Between March 20*[ti] OR Between April 20*[ti] OR Between May 20*[ti] OR Between June 20*[ti] OR Between July 20*[ti] OR Between August 20*[ti] OR Between September 20*[ti] OR Between October 20*[ti] OR Between November 20*[ti] OR Between December 20*[ti] OR between 19*[ti] OR between 20*[ti]) OR Meta-Analysis[ptyp] OR "Diffusion of Innovation"[mh] OR Cohort Studies[mh])

NOT ("Letter" [ptyp] OR ("Clinical Trial" [Publication Type] NOT "Clinical Trials as Topic"[mh]) OR ("animals"[MeSH Terms] NOT "humans"[MeSH Terms]))

OR drugs@fda[tiab]

Second search

Search date: 07 January 2018: 11856 hits. Search terms:

("Drug Approval"[Mesh] OR (approval[All Fields] OR approval'[All Fields] OR approval''[All Fields] OR approval's[All Fields] OR approvalall[All Fields] OR approvalethical[All Fields] OR approvalfor[All Fields] OR approvalformal[All Fields] OR approvalof[All Fields] OR approvalour[All Fields] OR approvalr[All Fields] OR approvals[All Fields] OR approvalsampling[All Fields] OR approvalsdata[All Fields] OR approvalsresearch[All Fields] OR approvalsthis[All Fields] OR approvalthe[All Fields] OR approvalthis[All Fields] OR approvalwe[All Fields] OR approvalwritten[All Fields]) OR ("licensure"[MeSH Terms] OR "licensure"[All Fields] OR "licensing"[All Fields]))

AND (("neoplasms"[MeSH Terms] OR "neoplasms"[All Fields] OR "cancer"[All Fields]) OR ("neoplasms"[MeSH Terms] OR "neoplasms"[All Fields] OR "oncology"[All Fields]) OR ("haematology"[All Fields] OR "hematology"[MeSH Terms] OR "hematology"[All Fields]))

5. Clinical trial evidence supporting US FDA approval of novel cancer therapies between 2000 and 2016

Aviv Ladanie [1, 2], Benjamin Speich [1], Florian Naudet [3], Arnav Agarwal [4, 5], Tiago V Pereira [6], Francesco Sclafani [7], Juan Martin-Liberal [8, 9], Thomas Schmid [10], Hannah Ewald [1, 2, 11], John PA Ioannidis [12, 13, 14, 15], Heiner C Bucher [1], Benjamin Kasenda [1, 17], Lars G Hemkens [1]

[1] Basel Institute for Clinical Epidemiology and Biostatistics, Department of Clinical Research, University Hospital and University of Basel, Basel, Switzerland
[2] Swiss Tropical and Public Health Institute (Swiss TPH), Basel, Switzerland
[3] Univ Rennes, CHU Rennes, Inserm, CIC 1414 [(Centre d'Investigation Clinique de Rennes)], F-35000 Rennes, France
[4] Department of Medicine, University of Toronto, Toronto, Ontario, Canada
[5] Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, Ontario, Canada
[6] Health Technology Assessment Unit, Institute of Education and Health Sciences, Oswaldo Cruz German Hospital, São Paulo, Brazil
[7] Department of Medicine, The Royal Marsden NHS Foundation Trust, Sutton, Surrey, UK
[8] Catalan Institute of Oncology (ICO) Hospitalet, Melanoma, Sarcoma and GU Tumors Unit, Barcelona, Spain
[9] Vall d'Hebron Institute of Oncology (VHIO), Early Drug Development Unit (UITM), Barcelona, Spain
[10] St. Clara Hospital, Basel, Switzerland
[11] University Medical Library, University of Basel, Basel, Switzerland
[12] Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, California, USA
[13] Department of Medicine, Stanford University School of Medicine, Stanford, California, USA
[14] Department of Health Research and Policy, Stanford University School of Medicine, Stanford, California, USA
[15] Department of Biomedical Data Science, Stanford University School of Medicine, Stanford, California USA
[16] Department of Statistics, Stanford University School of Humanities and Sciences, Stanford, California, USA
[17] Medical Oncology, University Hospital and University of Basel, Basel, Switzerland

Status: This manuscript is in draft form. The results were interpreted by Aviv Ladanie and Lars G Hemkens for this thesis with input from several co-authors. The current draft has not been approved for publication by the other authors.

Abstract

Importance: Clinical trial evidence supporting US Food and Drug Administration (FDA) approval of novel cancer drugs is typically the only information on benefits and harms available to patients and clinicians for decision-making when novel cancer therapies become available. Various evaluations raised concerns about the uncertainty surrounding the validity of this evidence.

Objective: To describe the clinical trial evidence available at US Food and Drug Administration approval of novel cancer drugs between 2000 to 2016.

Design and Setting: Meta-epidemiological analysis of clinical trials that provide the key efficacy data reported in FDA drug approval packages obtained from drugs@FDA.

Main outcomes and measures: We described indication regulatory, and trial design characteristics, calculated summary treatment effects on overall survival (OS), progression-free survival (PFS) and objective tumor response (TR) across all therapies in random effects meta-analyses; and estimated median absolute survival gains in months. We separately assessed disease and regulatory characteristics in stratified analyses.

Results: 92 novel drugs were approved in 100 indications based on 127 clinical trials. The 127 clinical trials included a median number of 193 patients (interquartile range [IQR]: 106, 448). Half of the clinical trials were randomized controlled (51%) and most were open-label (75%). The hazard ratio (HR) for the pooled average treatment effect on OS was 0.77 (95% confidence interval [CI]: 0.73, 0.81; I-squared [I^2] = 47%) and 0.52 (95% CI: 0.47, 0.57; I^2 = 88%) on PFS. The odds ratio for TR was 3.58 (95% CI: 2.77, 4.62; I^2 = 87%). The median absolute survival gain was 2.40 months (IQR: 1.25, 3.89). There was moderate to substantial heterogeneity across individual treatment effects, but the overall range of effect sizes was widely similar across the various disease and regulatory strata.

Conclusions: Approval data from over 17 years indicates that patients and clinicians have typically limited evidence on benefits and harms when a novel cancer treatment enters the market. This evidence shows that novel cancer therapies typically have substantial effects on tumor response

endpoints, but they prolong the patient's overall survival on average only by about 2.40 months.

Keywords: US Food and Drug Administration (FDA), Drug regulation, Marketing authorization, Approval package, Drugs and biologics, Clinical trials, Cancer.

Introduction

Cancer research is characterized by the urgency to find novel drugs that improve patients' survival or health-related quality of life. Before patients get access to novel therapies, the available evidence on benefits and harms from clinical trials is assessed by authorities such as the United States Food and Drug Administration (FDA). Several regulatory programs have been established in the United States to expedite the development and review of drugs for the treatment of serious conditions (such as cancer) and rare diseases in general (1). This may allow patients to gain earlier access to novel drugs, but there is also concern that these programs may increase uncertainty in clinical decision-making as approvals under these regulations are often based on evidence from fewer and smaller trials (2). Furthermore, the clinical trials may be more prone to bias due to lack of randomization. However, previous analyses of approval evidence (3-6) are rather fragmented and either did not include the newer cancer treatments approved after 2013 (e.g., those with the novel breakthrough designation established in 2012), or evaluated no treatment effects, or were restricted to certain types of cancer (e.g., solid tumors) only.

We systematically investigated the available evidence for all novel cancer drugs approved by the FDA in the 17-year period of 2000 to 2016. We describe the regulatory characteristics and supporting clinical trials and determined the treatment effects on overall survival (OS), progression-free survival (PFS), and objective tumor response (TR).

Methods

This study is part of the Comparative Effectiveness of Innovative Treatments in Cancer project (CEIT-Cancer). We provide full details about the database and the data identification, selection, extraction, and handling process elsewhere (7, 8).

In brief, we identified all new drugs and therapeutic biologics approved by the FDA between the years 2000 and 2016 for the first time (i.e., as new molecular or biologic entities) for the treatment of solid tumors or hematological malignancies. We excluded supportive care or imaging drugs without anticancer activity. We obtained the corresponding FDA approval documents from the publicly available drugs@FDA database (9) and searched in the documents for randomized controlled (RCT) and single-arm trials (SAT; including randomized dose-comparison trials) that may have provided evidence on the benefits and harms of the novel treatment. We included all trials that were explicitly labeled "pivotal" to FDA approval decisions. In applications where no trial was described as explicitly pivotal, we selected a) RCTs enrolling the target population and comparing the novel drug with a control intervention (that did not contain the novel drug) and b) trials we deemed essential for approval (for example when extensively discussed in the medical review). We retrieved key design characteristics of trials as well as the reported treatment effects on OS, PFS, and TR. All steps and extractions were conducted by two reviewers working independently (AL and either AA, BK, BS, FN, FS, HE, JML, TS, or TVP). Any disagreement was resolved by consensus. Only the number of participants, and information on trial phase and type of blinding was extracted by one reviewer alone (AL).

Data analysis

We used descriptive statistics to analyze drugs, indications, and regulatory characteristics, key trial design features and the reported treatment effects. We used only RCTs for the analysis of treatment effects because non-randomized evidence is typically of limited value to determine benefits and harms of treatments. In five three-arm RCTs evaluating the experimental treatment in two different doses, we

selected the treatment comparison with the dose that was later approved.

We combined treatment effects from all RCTs in meta-analyses using random-effects models (10) and describe the statistical heterogeneity using the I-squared (I^2) statistic (11). Tumor response is presented here as crude (unadjusted) odds ratios (OR). Analyses were conducted overall and stratified by cancer type (solid tumors or hematologic malignancies), orphan status (with or without), and approval pathway (traditional or accelerated approval). All analyses are exploratory. We used Microsoft Excel, R Version 3.5.0/RStudio 1.1.383, and Stata 14.2. All averages are reported as medians.

Results

We identified 92 novel drugs approved between 2000 and 2016 (Table 1). Forty percent (40%, 37 drugs) were classified by the FDA as first-in-class (new mechanism of action), 40% (37) as advance-in-class (significant improvements over first-in-class drugs) and 12% (13) as addition-to-class (Table 1).

The 92 drugs were approved for 100 indications (7 drugs were approved for multiple indications) (Table 2), 58% as treatments for solid tumors, mostly of the genitourinary system (e.g., renal, prostate, or urothelial cancers) and 42% for hematologic malignancies, mostly leukemias. The majority were first line (36%) and second line (45%) treatments, and the remaining 19% were third or fourth line treatments. The accelerated approval program was used for 44% of the indications, and 66% had an orphan status. All treatments for hematological indications had an orphan status (Figure 1).

Characteristics of clinical trials

We included a total of 127 clinical trials with a median of 193 enrolled patients (interquartile range [IQR]: 106, 448) (Table 3). The median number of eligible and analyzed clinical trials per approved indication was 1 (IQR: 1, 2). The trials were larger when supporting solid, non-orphan, and non-accelerated approval indications compared to approvals supporting drugs for hematological malignancies, orphan indications, and accelerated approved indications (Table 4).

About half of all trials were RCTs (65 of 127, 51%) and phase 2 trials (66 of 127, 52%). Twenty-five percent (25%, or 30 of 127) were double-blinded. While approvals of treatments for solid tumors, non-orphan indication, and non-accelerated approval pathways were typically supported by RCTs (51 of 72, 71%; 31 of 41, 76%; and 54 of 66, 82%), SATs were the typical study design for hematological malignancies (41 of 55; 75%), orphan indications (52 of 86; 60%) and accelerated approved indications (50 of 61; 82%). The same pattern was observed for trial phase. Solid tumor, non-orphan and non-accelerated approved indications were typically supported by phase 3 trials (61%, 66%, and 76%). Malignant hematology, orphan, and accelerated approved indications were typically supported by phase 2 trials (75%, 60%, and 83%) (Table 4). Trials for solid tumors, non-orphan and non-accelerated approvals were sometimes double-blinded (35%, 39%, and 41%), while this was rare in trials for hematological malignancies, for orphan diseases, and accelerated approval pathways (9%, 16%, and 5%).

Overall, half of all trials had no parallel control without the experimental drug (historical control 42% [54 of 127] and dose-comparison 6% [8 of 127]). Trials supporting approval of drugs for solid tumors were predominantly active-controlled (46%). The same was the case for drugs without an orphan status (51%) as well as drugs that received non-accelerated or traditional approval (52%). Drugs approved for hematological malignancies, orphan diseases, as well as accelerated approved drugs typically had no parallel control group (Table 3; 69%; 56%; 70%, respectively).

Treatment effects

Across all 54 RCTs with reported treatment effects on OS, the combined risk to die from any cause across all novel cancer treatments was reduced on average by 23% (hazard ratio [HR] = 0.77; 95% confidence interval [CI]: 0.73, 0.81; I^2 = 47%; Figure 2, Table 4) with a median survival gain of 2.40 months (IQR: 1.25, 3.89; Table 4).

Across all treatments, the combined risk for tumor progression or death (PFS) from any cause was reduced on average by 48% (HR = 0.52; 95% CI: 0.47, 0.57; I^2 = 88%; Figure 3, Table 4) with a median

PFS gain of 2.72 months (IQR: 1.61, 4.45; Table 4). Patients receiving the novel treatment had on average a 3.58-fold greater tumor response (OR = 3.58; 95% CI: 2.77, 4.62; I^2 = 87% Figure 4; Table 4).

Treatment effects on OS are remarkably consistent across subsets for solid tumor indications (HR = 0.76, 95% CI: 0.72, 0.80), orphan and non-orphan indications (HR = 0.77, 95% CI: 0.71, 0.85; HR 0.76, 95% CI: 0.72, 0.81), and indications granted full (non-accelerated) approval (HR = 0.76, 95% CI: 0.72, 0.80) (Figures 5-7; Table 4). Treatment effects for hematological malignancies and accelerated approvals were slightly smaller (HR = 0.87, 95% CI: 0.76, 0.99; HR = 0.85, 95% CI: 0.71, 1.01). Overall, median OS gains were similar across all subsets, ranging from 2.15 months (for hematological malignancies) to 3.20 months (accelerated approval).

Treatment effects on PFS are also consistent across subsets for solid tumor indications (HR = 0.53, 95% CI: 0.48, 0.58), non-orphan indications (HR = 0.59, 95% CI: 0.54, 0.66), and accelerated and full (non-accelerated) approvals (HR = 0.54, 0.44, 0.67; HR 0.51, 95% CI: 0.46, 0.57) (Figures 8-10; Table 4). Treatment effects for hematological malignancies and orphan indications were slightly stronger (HR = 0.43, 95% CI: 0.28, 0.66; HR = 0.44, 95% CI: 0.38, 0.58). Overall, median PFS gains were similar across all subsets, ranging from 2.08 months (for non-orphan indications) to 4.30 months (hematological malignancies).

Treatment effects on tumor response ranged from 2.61-fold for solid tumor indications (OR = 2.61, 95% CI: 1.88, 3.62) to 4.71-fold for orphan indications (OR = 4.71, 95% CI: 3.24, 6.84) (Figures 11-13; Table 4).

The statistical heterogeneity of OS effect sizes between the trials was moderate (I^2 = 47%, Figure 2) and was high for PFS (I^2 = 88%, Figure 3) and TR (I^2 = 87%, Figure 4). This was not explained by disease type, orphan status or approval type.

Discussion

Over 17 years, 92 novel therapies for various types of cancer were approved by the US FDA on the basis of 127 clinical trials. The typical cancer drug was approved on the basis of one trial meeting our inclusion criteria. Overall, the trials were mostly non-blinded, half of them lacked randomization and half of them included fewer than 200 patients. This evidence available at approval shows that novel drug treatments typically have substantial effects on tumor response, but they prolong the patient's overall survival on average only by about 2.40 months.

Drugs approved as treatment for hematologic malignancies either have an orphan status, underwent accelerated approval pathways, or both. The evidence supporting their approval typically comes from a relatively small number of patients, predominantly enrolled in single-arm and phase 2 trials. In contrast, drugs for solid tumors, without orphan indications, or without accelerated approval pathways entered the market with evidence more frequently from randomized trials and larger patient groups.

Relative treatment effects were stronger on surrogate outcomes (PFS and TR) than on overall survival. There was moderate to high statistical heterogeneity across treatment effect sizes, but the overall range of effect sizes was widely similar, also across the various subsets. The only drug that showed absolute survival gains of more than 6 months compared to the control group was LARTRUVO® (olaratumab), which had an orphan status and was granted accelerated approval in 2016 as sarcoma therapy. The pivotal trial demonstrated a relatively large survival gain of almost 1 year, while PFS was prolonged by only 2.5 months without nominal statistical significance. This exemplifies the relevance of careful considerations of surrogacy issues in benefit assessments (12, 13).

Our findings are broadly consistent with other evaluations of approval evidence from the US and Europe. Downing et al. (2) analyzed all novel drugs and therapeutic biologics approved by the FDA between 2005 and 2012 for the first time. They estimated the median number of patients enrolled in cancer trials to be 266 (IQR: 84, 610). Less than half of cancer trials were randomized controlled (22/55, 47%), and only 27% (15/55) were double-blinded. Kesselheim et al. (14) analyzed all novel cancer drugs

that received FDA approval between 2004 and 2010 for orphan indications. The 23 pivotal trials had a median enrolment of 96 patients (IQR: 66, 152) and only 7 (30%) were randomized controlled.

Fojo et al. (15) found for drug approvals for solid tumor indications granted by the FDA between 2002 and 2014 a median survival gain of 2.10 months for OS and 2.5 months for PFS. Salas-Vega et al. (16) analyzed all novel cancer drugs approved by the FDA and EMA between 2003 and 2013 and estimated a mean OS gain of 3.43 months overall and 2.61 months across all malignant hematology indications. Davis et al. (3) recently reported a median OS gain of 2.70 months (range 1.00 to 5.80 months) across 48 cancer drugs approved by the EMA between 2009 and 2013 in 68 indications. This collective evidence, including our findings, indicates that even if a survival gain would always exist, it would be rather small in most cases.

This study has some limitations. First, our analysis was restricted to evidence presented to the FDA and reported in the approval packages. There may be other studies evaluating the benefits and harms of the drugs in these indications and we have not searched for them. We would assume that a manufacturer would present the best and most rigorous supporting evidence to the FDA to get approval. Moreover, our sample not only includes trials that were explicitly labeled as pivotal, we included any randomized trial in the same target population which may provide further evidence. Thus, we do not believe there is any publication bias in our sample that would lead to an underestimation of the utility of the evaluated therapies.

Second, our analysis was restricted to drugs that were finally approved and entered the market. This is a strongly selected sample and as such prone to substantial regression-to-the mean effects (17). This would lead to inflated benefits and again to an overestimation of the utility of the evaluated therapies.

Third, we focused on drugs indications, clinical and regulatory details, trial characteristics and the reported treatment effects but conducted no appraisal of the evidence with a thorough risk of bias assessment. While major sources of bias resulting from lack of blinding and lack of randomization are reflected in our data, we undertook no detailed assessment of each study. For example, we did not evaluate if the outcome assessment in the open label trials was blinded, if the duration of treatments

and follow-ups were adequate, or if the control groups were optimal. The latter is an important limitation as suboptimal controls would lead to overestimated benefits of the experimental drug. Again, this would have led us overestimate the benefits of the evaluated drugs.

Fourth, while we evaluated effects on OS, PFS, and TR because they are frequently used in cancer trials, we did not collect data on quality of life and side effects. Given the comparatively small survival gains associated with the use of these novel treatments, the quality of these last days of patients' life is certainly of utmost importance and deserves closer attention in future research.

Overall, the data from over 17 years of novel cancer drug approvals indicates that patients and clinicians have typically limited evidence on the benefits and harms when a novel cancer treatment enters the market – in only half of indications, data from an RCT are available. This is even more problematic for patients with hematological malignancies. While thus far these novel therapies may have substantial effects on tumor size or other markers of tumor response, on average, they prolong the life of patients by 2.40 months. Many of these drugs were approved to address an unmet medical need. We believe that this need still exists.

Tables and figures

TABLE 1: DRUG CHARACTERISTICS

	(N = 92) n (%)
Type of substance	
Chemical	67 (73)
Biologic	25 (27)
Innovation class	
First-in-class	37 (40)
Advance-in-class	37 (40)
Addition-to-class	12 (13)
Unknown *	6 (7)

* Innovation status of n = 6 drugs could not be ascertained.

TABLE 2: INDICATION CHARACTERISTICS

	(N = 100) n (%)
Treatment	
Cancer type	
Solid tumors	58 (58)
Genitourinary	14 (14)
Digestive/Gastrointestinal	10 (10)
Respiratory/Thoracic	9 (9)
Skin	9 (9)
Breast	8 (8)
Endocrine and Neuroendocrine	3 (3)
Gynecologic	2 (2)
Musculoskeletal	2 (2)
Neurologic	1 (1)
Hematologic malignancies	42 (42)
Leukemia	18 (18)
Lymphoma	12 (12)
Multiple myeloma	7 (7)
Other *	5 (5)
Line of treatment	
1st	36 (36)
2nd	45 (45)
3rd	15 (15)
4th	4 (4)
Regulatory	
Priority review designation **	77 (77)
Approved with an orphan designation	66 (66)
Approved under accelerated approval	44 (44)
Breakthrough therapy designation approval ***	15 (15)

* Encompasses myelodysplastic syndromes (n = 2), acute lymphoblastic leukemia/lymphoma, myelofibrosis and multicentric Castleman's disease; ** Priority review designation status of four indications could not be ascertained. *** Beginning for approvals in 2013 and later.

TABLE 3: PIVOTAL TRIAL CHARACTERISTICS

		Cancer type		Orphan status		Accelerated approval	
	Overall	Solid tumors	Hematologic malignancies	No	Yes	No	Yes
N subjects enrolled, median (IQR)	193 (106, 448)	330 (171, 638)	111 (74, 188)	435 (230, 760)	152 (94, 292)	374 (159, 710)	136 (100, 202)
	N = 127, n (%)	N = 72, n (%)	N = 55, n (%)	N = 41, n (%)	N = 86, n (%)	N = 66, n (%)	N = 61, n (%)
Study design							
Randomized controlled trial	65 (51)	51 (71)	14 (25)	31 (76)	34 (40)	54 (82)	11 (18)
Single-arm trial *	62 (49)	21 (29)	41 (75)	10 (24)	52 (60)	12 (18)	50 (82)
Trial phase (drug development phase)							
Phase 3	57 (45)	44 (61)	13 (24)	27 (66)	30 (35)	50 (76)	7 (11)
Phase 2	66 (52)	25 (35)	41 (75)	14 (34)	52 (60)	15 (23)	51 (83)
Phase 1	3 (2)	3 (4)	0 (0)	0 (0)	3 (3)	0 (0)	3 (5)
Not reported	1 (1)	0 (0)	1 (2)	0 (0)	1 (1)	1 (2)	0 (0)
Type of blinding employed							
Double-blind	30 (24)	25 (35)	5 (9)	16 (39)	14 (16)	27 (41)	3 (5)
Single-blind	1 (1)	1 (1)	0 (0)	0 (0)	1 (1)	1 (2)	0 (0)
Open label	95 (75)	45 (62)	50 (91)	25 (61)	70 (81)	37 (56)	58 (95)
Blinded (not specified)	1 (1)	1 (1)	0 (0)	0 (0)	1 (1)	1 (2)	0 (0)
Type of control group							
Parallel, without experimental drug	65 (51)	51 (71)	14 (25)	31 (76)	34 (40)	54 (82)	11 (18)
(Randomized) active **	43 (34)	33 (46)	10 (18)	21 (51)	22 (26)	34 (52)	9 (15)
(Randomized) placebo	19 (15)	17 (24)	2 (4)	9 (22)	10 (12)	18 (27)	1 (2)
(Randomized) no treatment ***	3 (2)	1 12)	2 (4)	1 (2)	2 (2)	2 (3)	1 (2)
No or other control group	62 (49)	21 (29)	41 (75)	10 (24)	52 (60)	12 (18)	50 (82)
(Randomized) dose-comparison	8 (6)	5 (7)	3 (5)	4 (10)	4 (5)	1 (2)	7 (11)
(Non-randomized) historical control	54 (42)	16 (22)	38 (69)	6 (15)	48 (56)	11 (17)	43 (70)

Type of endpoint with effects (RCTs only)	N = 65, n (%)	N = 51, n (%)	N = 14, n (%)	N = 31, n (%)	N = 34, n (%)	N = 54, n (%)	N = 11, n (%)
Overall survival	54 (83)	44 (86)	10 (71)	25 (81)	29 (85)	48 (89)	6 (55)
Progression-free survival	54 (83)	46 (90)	8 (57)	26 (84)	28 (82)	47 (87)	7 (64)
Objective tumor response	51 (78)	40 (78)	77 (79)	24 (77)	27 (79)	45 (83)	6 (55)

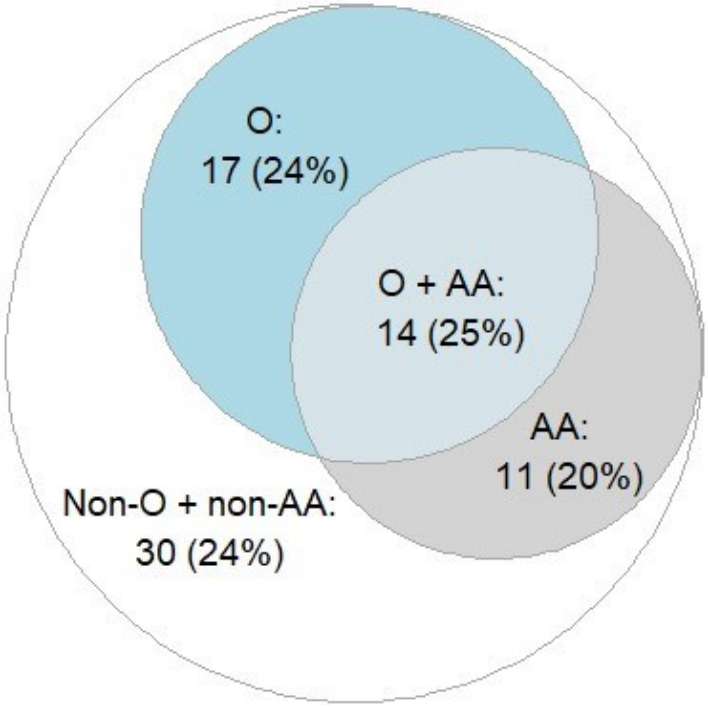
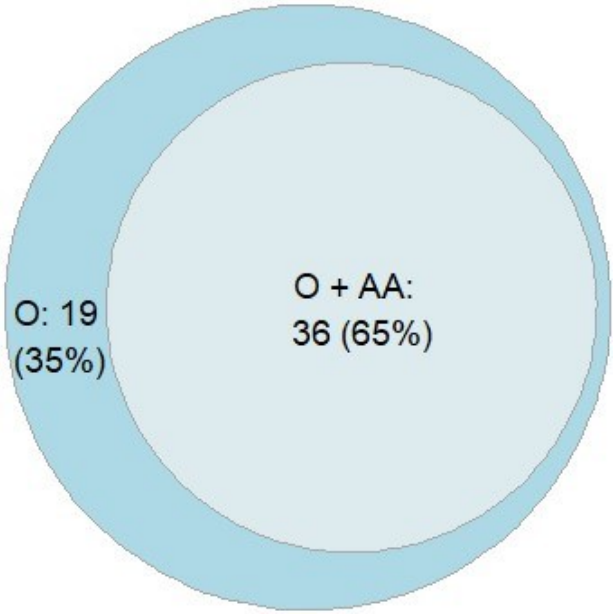
* Trials without a parallel control group, i.e., patients are randomized to different doses of the experimental treatment only (dose-comparison trials) or not randomized and compared to historical controls; ** Includes also comparators where placebo is given on top of an active treatment (add-on or double-dummy trials); *** Encompasses supportive therapy or usual care. Abbreviations: IQR, interquartile range.

TABLE 4: TREATMENT EFFECTS ON OVERALL SURVIVAL, PROGRESSION-FREE SURVIVAL, AND OBJECTIVE TUMOR RESPONSE

		Overall		Cancer type				Orphan status				Accelerated approval			
				Solid tumors		Hematologic malignancies		No		Yes		No		Yes	
		N RCTs	Effect	N RCTs	Effect	N RCTs	Effect	N RCTs	Effect	N RCTs	Effect	N RCTs	Effect	N RCTs	Effect
Overall survival															
Summary (95% CI) *	HR	54	0.77 (0.73, 0.81)	44	0.76 (0.72, 0.80)	10	0.87 (0.76, 0.99)	25	0.76 (0.72, 0.81)	29	0.77 (0.71, 0.85)	48	0.76 (0.72, 0.80)	6	0.85 (0.71, 1.01)
Median gains in months (IQR)		54	2.40 (1.25, 3.89)	44	2.40 (1.40, 3.89)	10	2.15 (0.61, 3.58)	25	2.40 (1.44, 4.10)	29	2.75 (0.72, 3.66)	48	2.22 (1.18, 3.66)	6	3.20 (2.70, 4.20)
Progression-free survival															
Summary (95% CI) *	HR	53**	0.52 (0.47, 0.57)	46	0.53 (0.48, 0.58)	7**	0.43 (0.28, 0.66)	26	0.59 (0.54, 0.66)	27**	0.44 (0.38, 0.53)	46**	0.51 (0.46, 0.57)	7	0.54 (0.44, 0.67)
Median gains in months (IQR)		54	2.72 (1.61, 4.45)	46	2.40 (1.48, 4.26)	8	4.30 (4.95, 10.1)	26	2.08 (1.10, 4.18)	28	3.50 (2.40, 4.64)	47	2.77 (1.68, 4.40)	7	2.50 (1.30, 5.50)
Objective tumor response															
Summary (95% CI) *	OR	50**	3.58 (2.77, 4.62)	40	3.48 (2.60, 4.65)	10	4.13 (2.30, 7.40)	24	2.61 (1.88, 3.62)	26**	4.71 (3.24, 6.84)	44**	3.68 (2.79, 4.85)	6	2.91 (1.40, 6.04)

* Random effects meta-analysis. ** The 95% confidence interval on PFS and TR endpoints was unavailable for one RCT. Abbreviations: CI, confidence interval; HR, hazard ratio; IQR, interquartile range; OR, odds ratio; PFS, progression-free survival; RCT, randomized controlled trial.

FIGURE 1: ORPHAN STATUS AND APPROVAL PATHWAY OF ALL NOVEL CANCER DRUGS BETWEEN 2000 AND 2016

Solid tumors (n = 72)	Hematologic malignancies (n = 55)
	

Venn diagram showing proportions of indications with orphan status and accelerated approval and their overlap. Color legend: blue, orphan indication (O); grey, accelerated approved indications (AA); white, standard approval without orphan indications. Left: Of 72 approvals for solid tumors, 30 were without orphan status or accelerated approval. 17 had orphan status alone, 11 received accelerated approval alone, and 14 had both an orphan status and were accelerated approved. Right: All 48 approvals for hematologic malignancies had orphan status, and the majority (36) also received accelerated approval. Abbreviations: AA, accelerated approval; O, orphan indication; Non-AA, traditional approval (not accelerated approval); Non-O, non-orphan indication.

FIGURE 2: FOREST PLOT OF ALL TRIALS WITH OVERALL SURVIVAL DATA

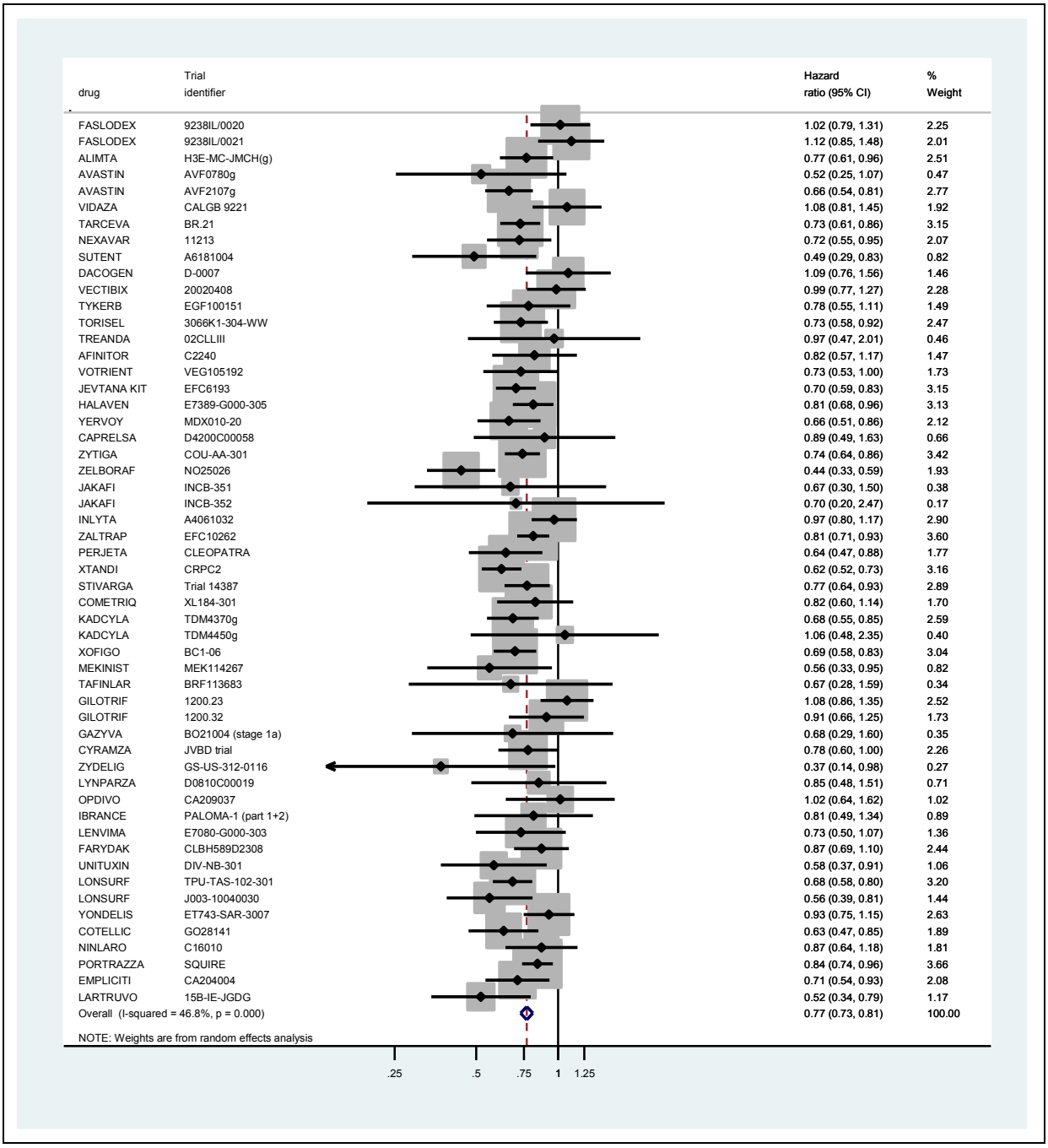


FIGURE 3: FOREST PLOTS OF ALL TRIALS WITH DATA ON PROGRESSION-FREE SURVIVAL

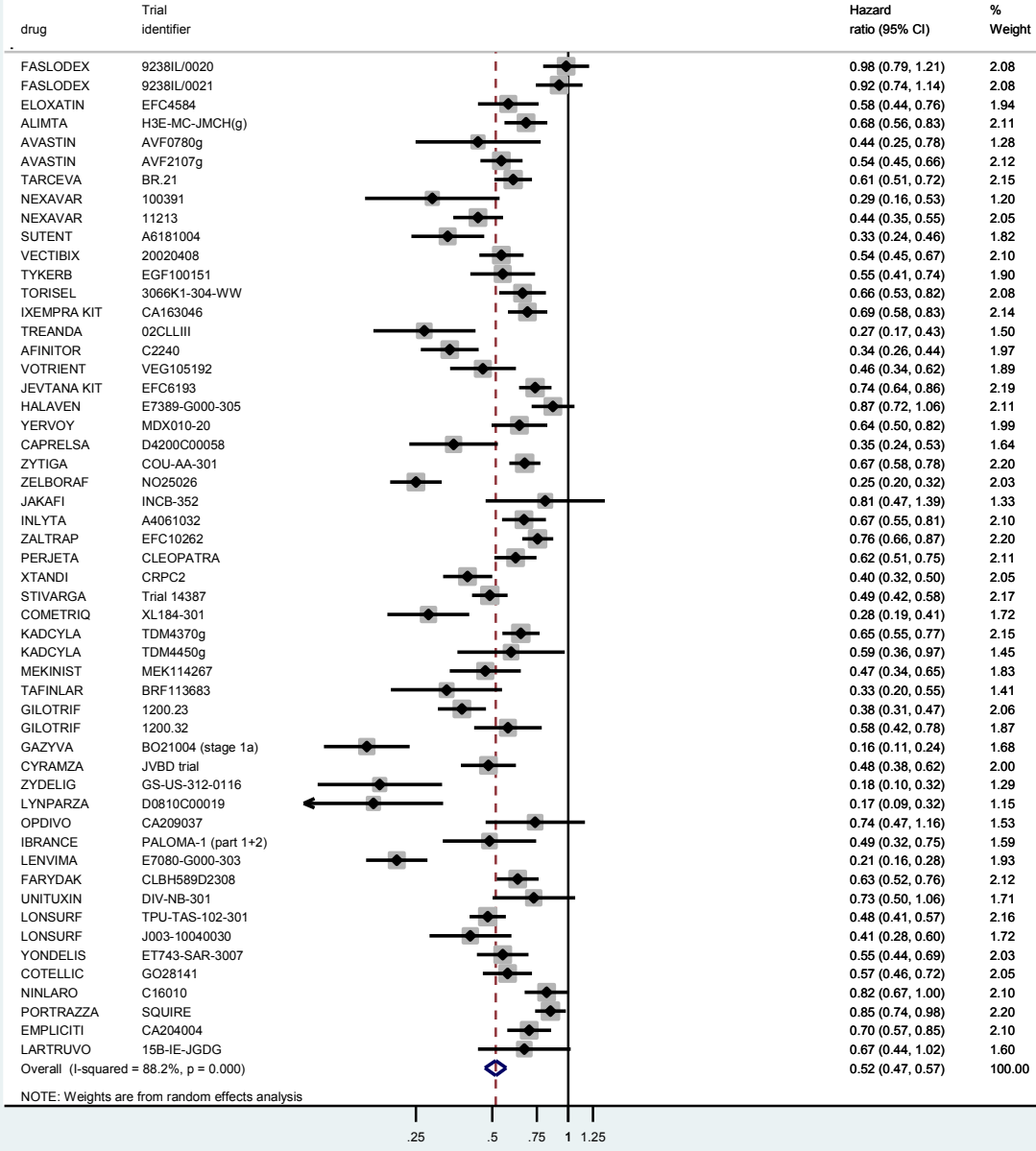


FIGURE 4: FOREST PLOTS OF ALL TRIALS WITH DATA ON OBJECTIVE TUMOR RESPONSE

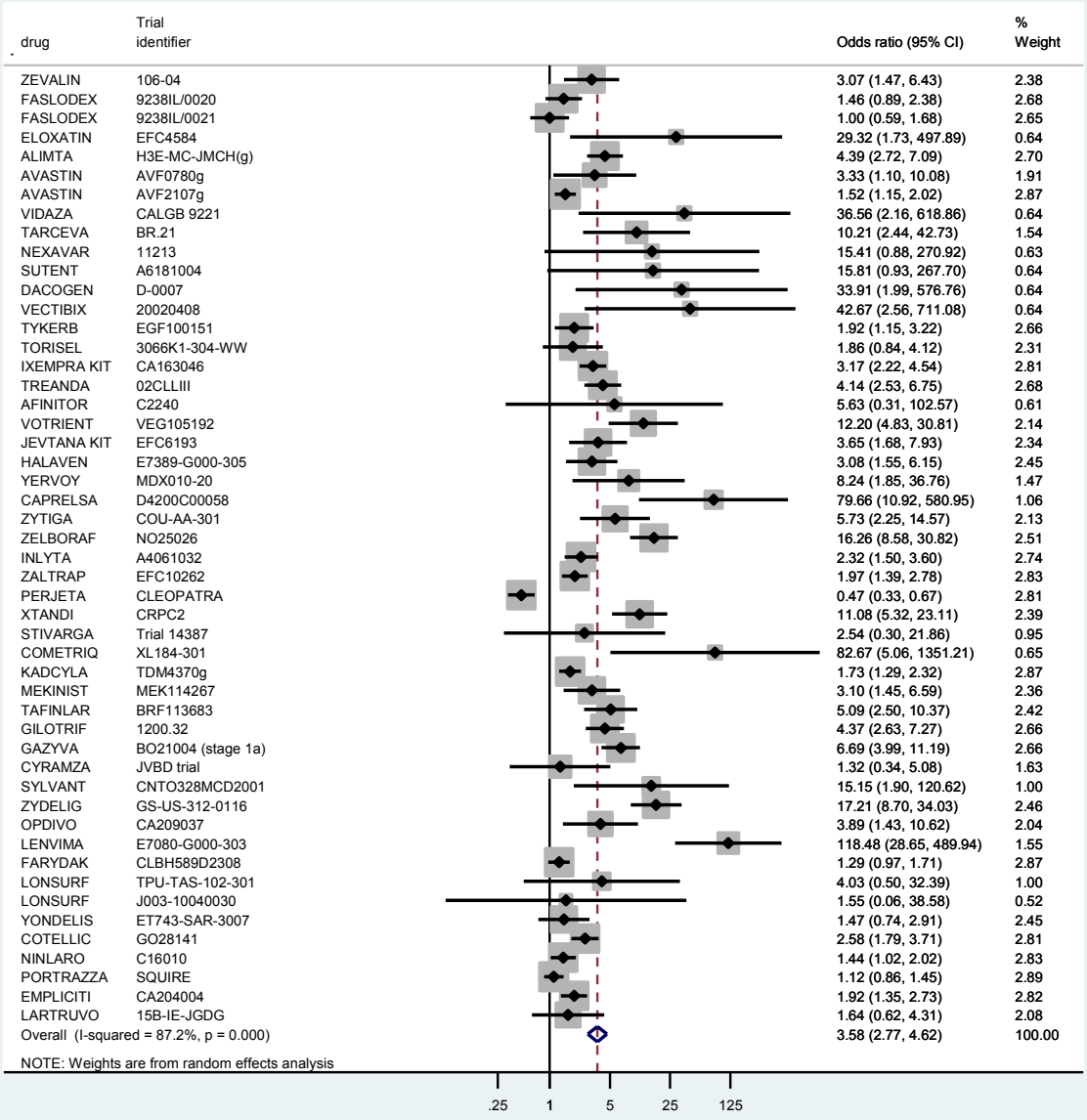


FIGURE 5: FOREST PLOTS OF ALL TRIALS WITH OVERALL SURVIVAL DATA, STRATIFIED BY DISEASE TYPE

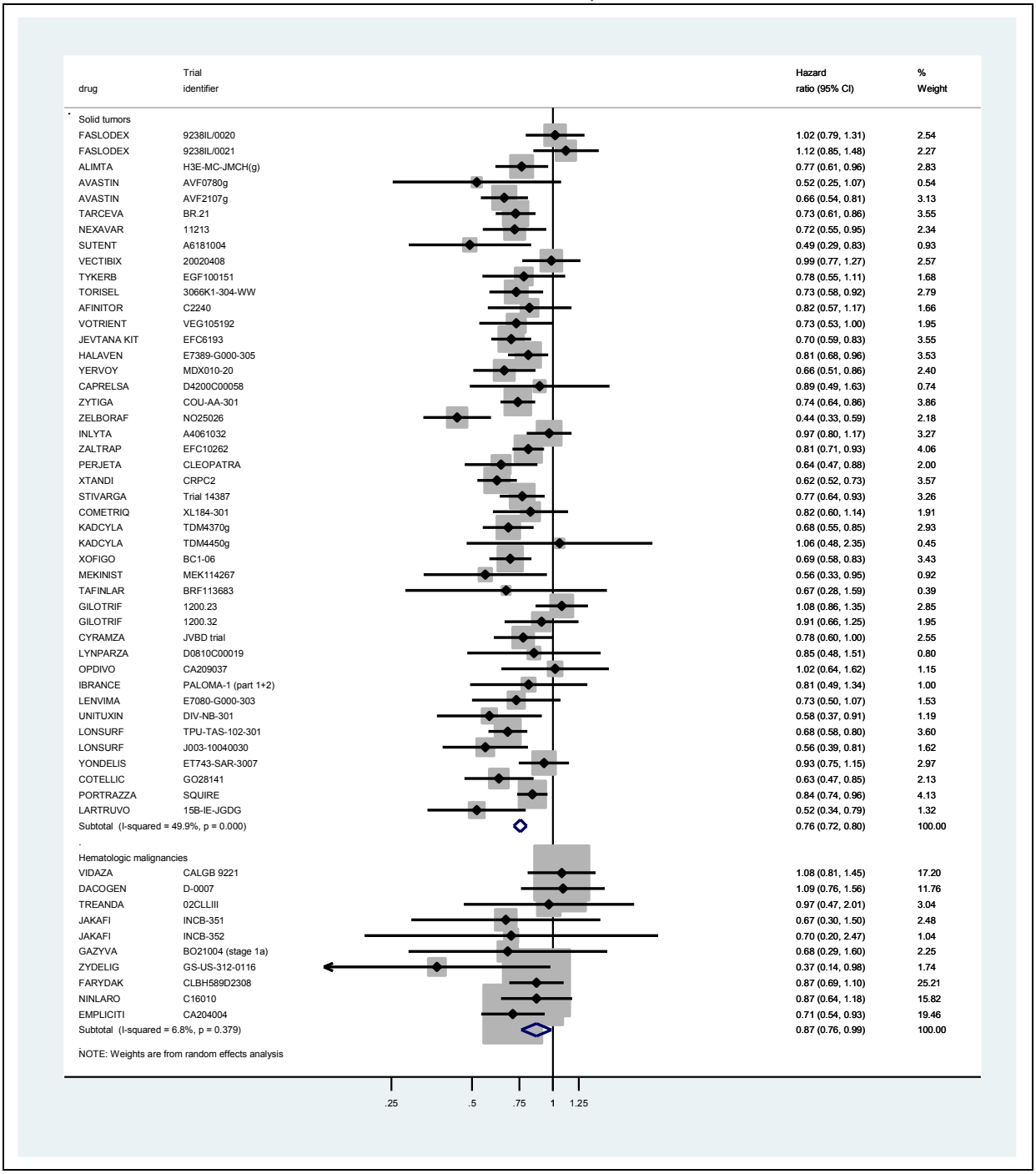


FIGURE 6: FOREST PLOTS OF ALL TRIALS WITH OVERALL SURVIVAL DATA, STRATIFIED BY ORPHAN STATUS

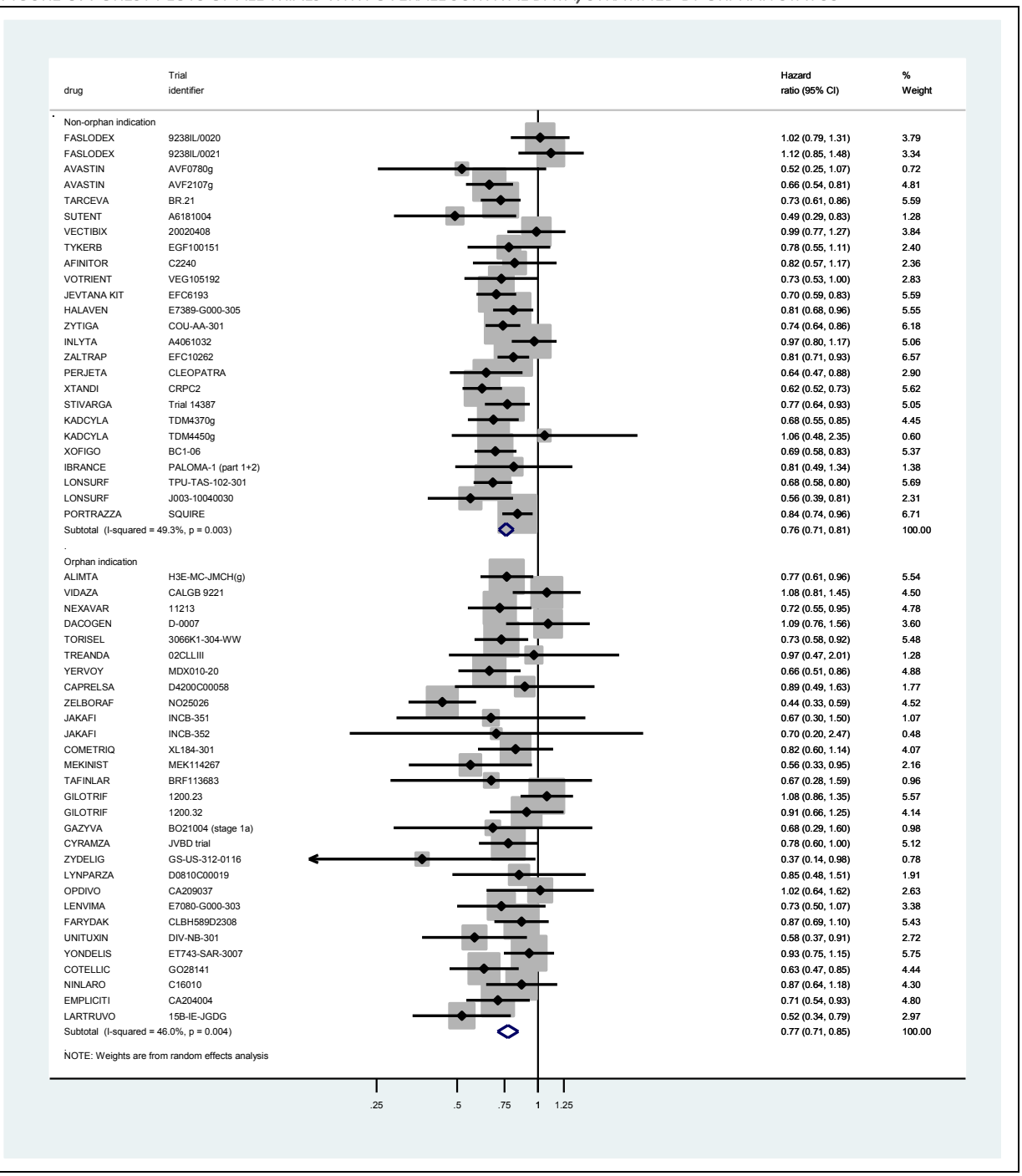


FIGURE 7: FOREST PLOTS OF ALL TRIALS WITH OVERALL SURVIVAL DATA, STRATIFIED BY APPROVAL TYPE

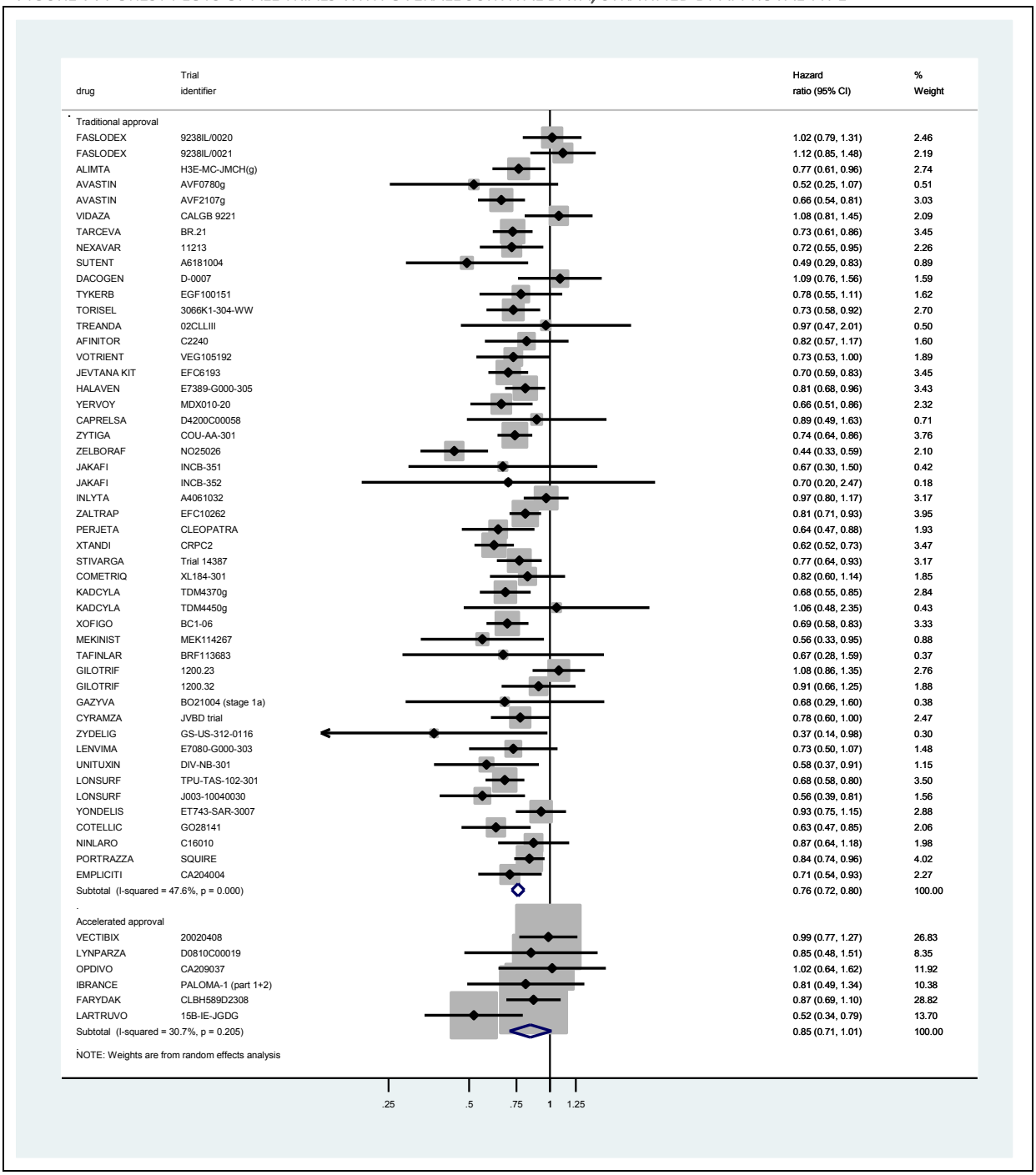


FIGURE 8: FOREST PLOTS OF ALL TRIALS WITH DATA ON PROGRESSION-FREE SURVIVAL, STRATIFIED BY DISEASE TYPE

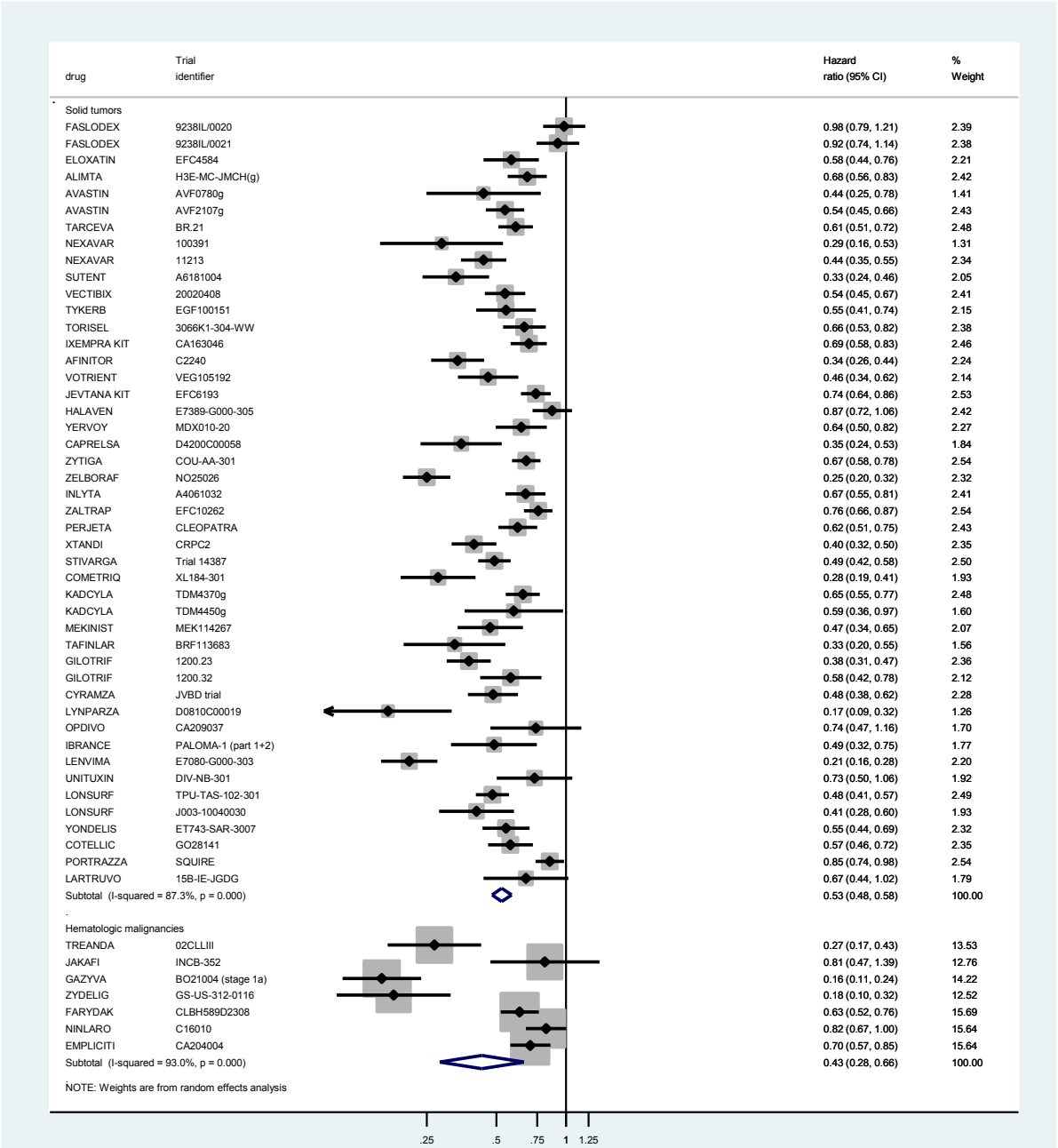


FIGURE 9: FOREST PLOTS OF ALL TRIALS WITH DATA ON PROGRESSION-FREE SURVIVAL, STRATIFIED BY ORPHAN STATUS

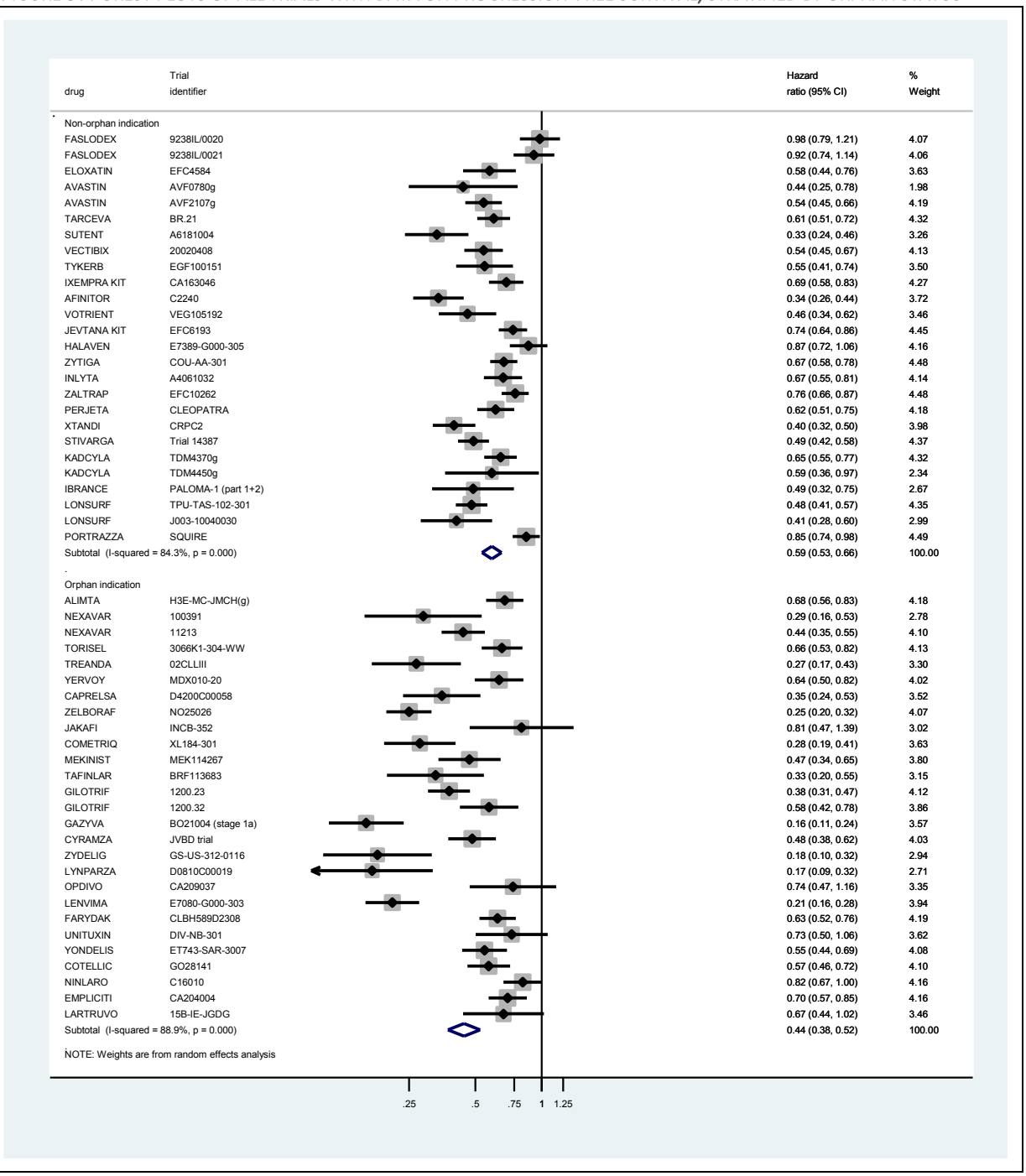


FIGURE 10: FOREST PLOTS OF ALL TRIALS WITH DATA ON PROGRESSION-FREE SURVIVAL, STRATIFIED BY APPROVAL TYPE

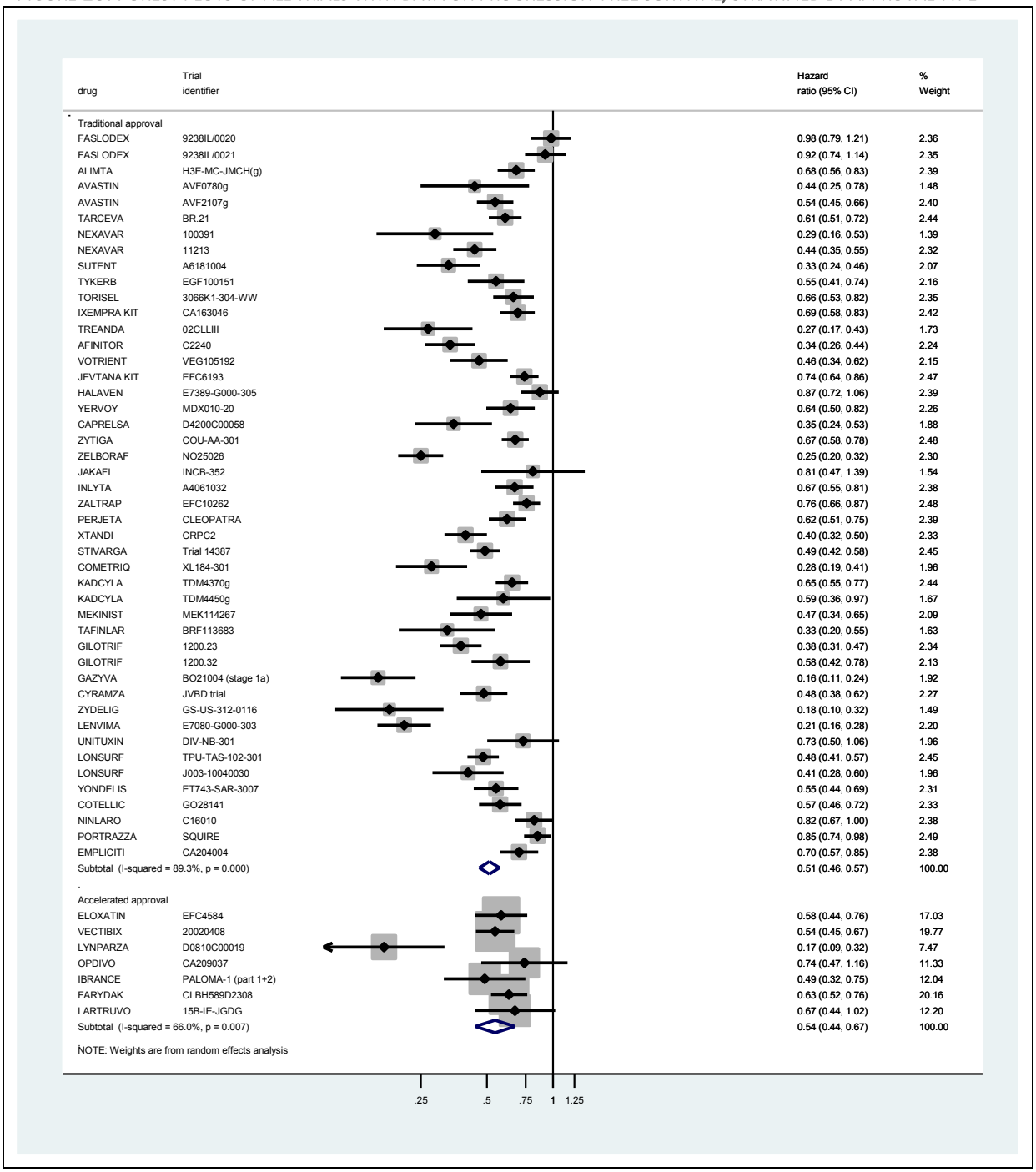


FIGURE 11: FOREST PLOTS OF ALL TRIALS WITH DATA ON OBJECTIVE TUMOR RESPONSE, STRATIFIED BY DISEASE TYPE

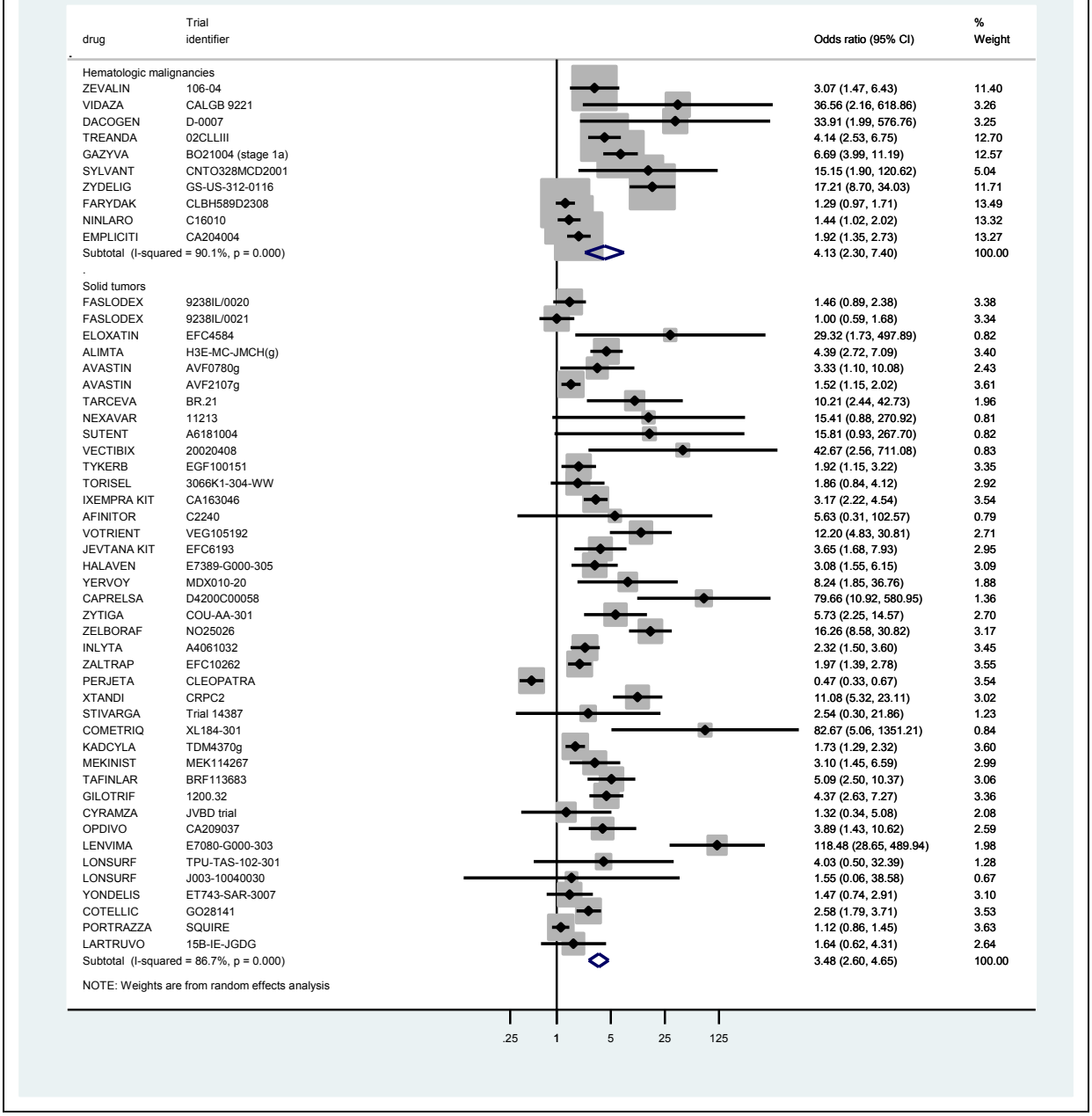


FIGURE 12: FOREST PLOTS OF ALL TRIALS WITH DATA ON OBJECTIVE TUMOR RESPONSE, STRATIFIED BY ORPHAN STATUS

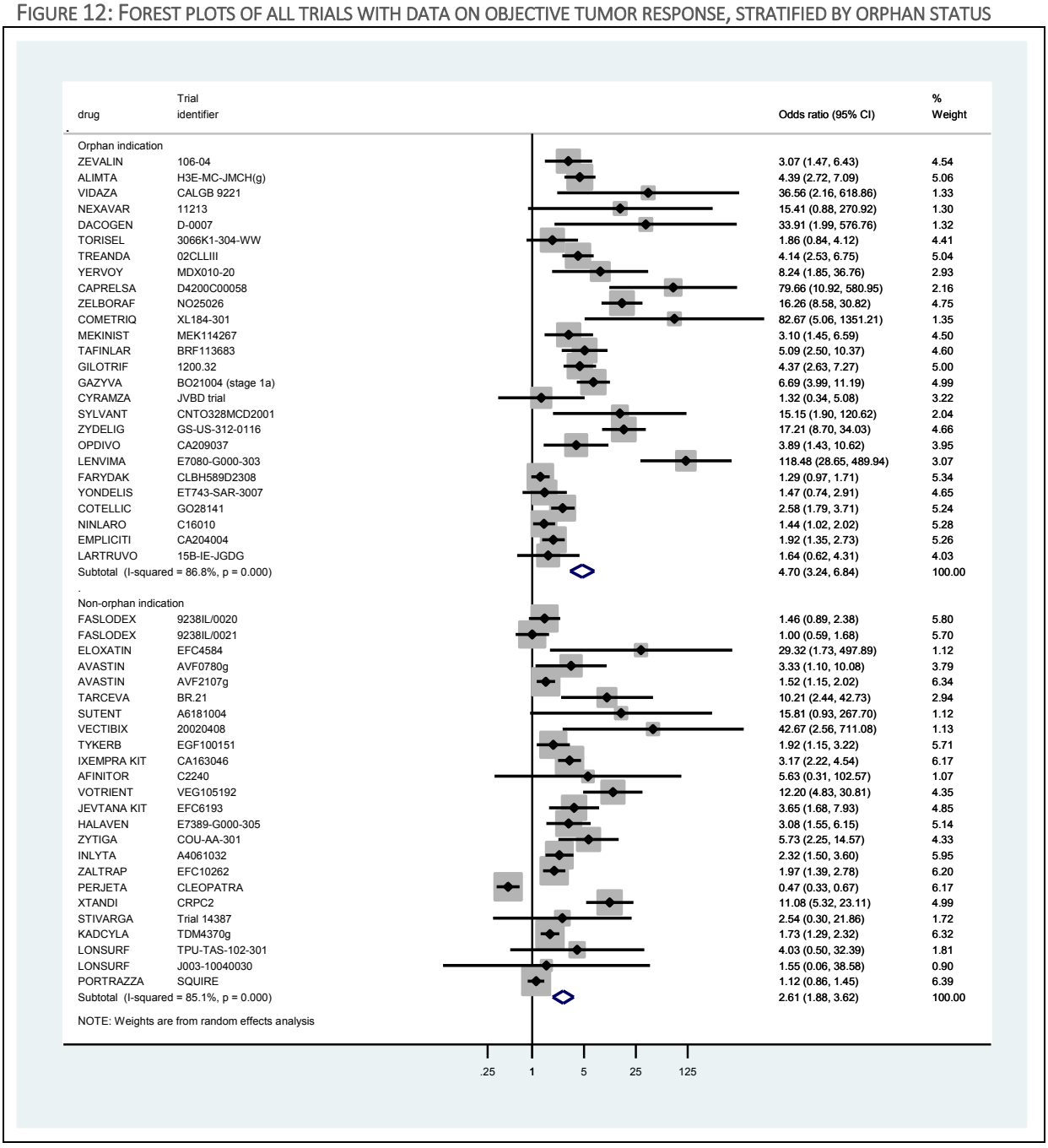
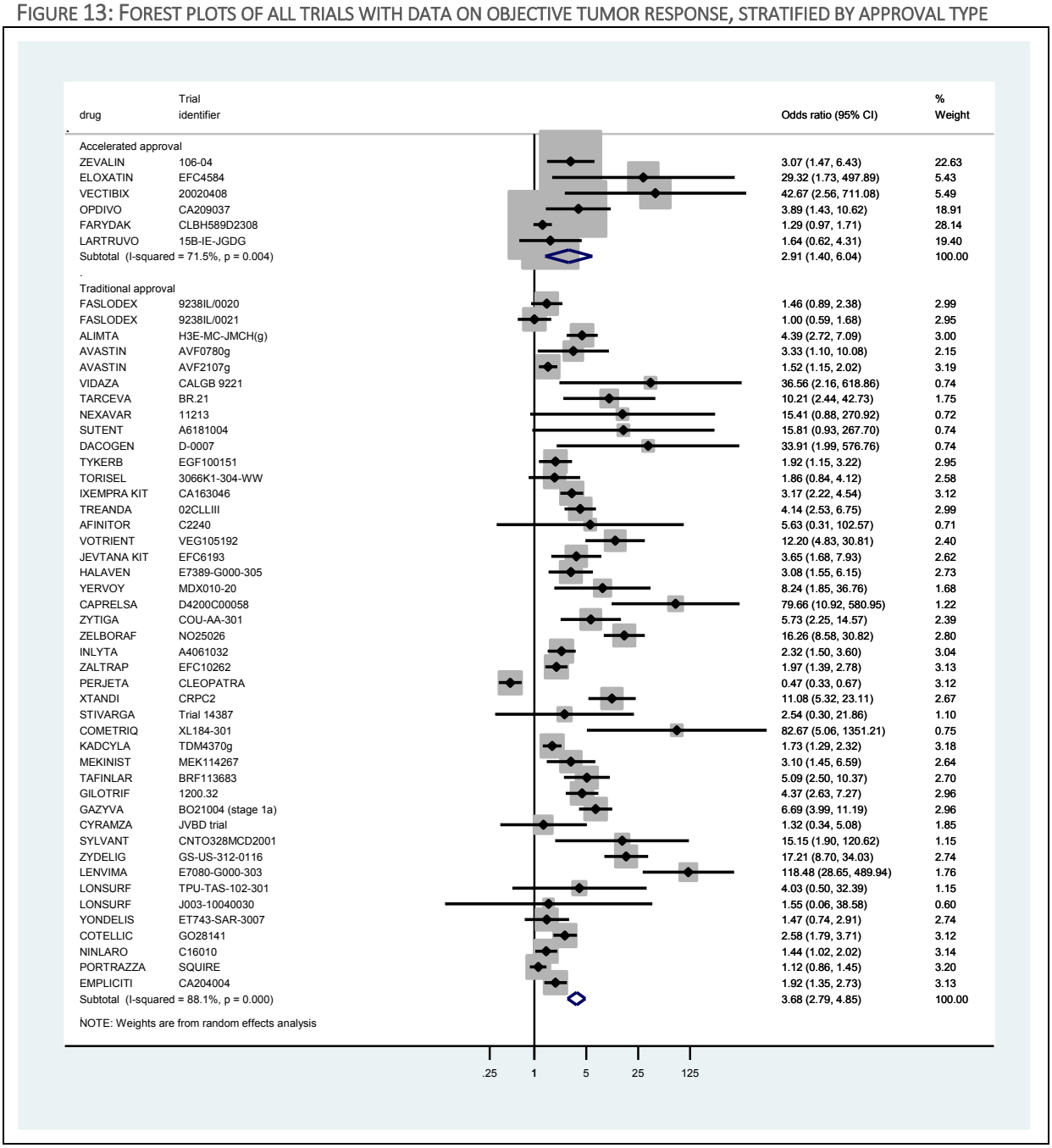


FIGURE 13: FOREST PLOTS OF ALL TRIALS WITH DATA ON OBJECTIVE TUMOR RESPONSE, STRATIFIED BY APPROVAL TYPE



References

1. US Food and Drug Administration. Guidance for industry: expedited programs for serious conditions – drugs and biologics. May 2014. Available from: www.fda.gov/downloads/Drugs/Guidances/UCM358301.pdf.
2. Downing NS, Aminawung JA, Shah ND, Krumholz HM, Ross JS. Clinical trial evidence supporting FDA approval of novel therapeutic agents, 2005-2012. JAMA 2014;311(4):368-77.
3. Davis C, Naci H, Gurpinar E, Poplavska E, Pinto A, Aggarwal A. Availability of evidence of benefits on overall survival and quality of life of cancer drugs approved by European Medicines Agency: retrospective cohort study of drug approvals 2009-13. BMJ 2017;359:j4530.
4. Hartmann M, Mayer-Nicolai C, Pfaff O. Approval probabilities and regulatory review patterns for anticancer drugs in the European Union. Crit Rev Oncol Hematol 2013;87(2):112-21.
5. Sridhara R, Johnson JR, Justice R, Keegan P, Chakravarty A, Pazdur R. Review of oncology and hematology drug product approvals at the US Food and Drug Administration between July 2005 and December 2007. J Natl Cancer Inst 2010;102(4):230-43.
6. Tsimberidou AM, Braiteh F, Stewart DJ, Kurzrock R. Ultimate fate of oncology drugs approved by the US Food and Drug Administration without a randomized trial. J Clin Oncol 2009;27(36):6243-50.
7. Ladanie A, Ewald H, Kasenda B, Hemkens LG. How to use FDA drug approval documents for evidence syntheses. BMJ 2018;362:k2815.
8. Ladanie A, Speich B, Naudet F, Agarwal A, Pereira TV, Sclafani F, et al. The Comparative Effectiveness of Innovative Treatments for Cancer (CEIT-Cancer) project: rationale and design of the database and the collection of evidence available at approval of novel drugs (manuscript accepted for publication in: Trials).
9. US Food and Drug Administration. Drugs@FDA: FDA approved drug products [Internet]. [cited 04 June 2018]. Available from: www.fda.gov/drugsatfda.

10. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7(3):177-88.

11. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327(7414):557-60.

12. Institute for Quality and Efficiency in Health Care. 2005. Validity of surrogate endpoints in oncology. Executive summary of rapid report A10-05, Version 1.1. Cologne, Germany.

13. Prasad V, Kim C, Burotto M, Vandross A. The strength of association between surrogate end points and survival in oncology: a systematic review of trial-level meta-analyses. JAMA Intern Med 2015;175(8):1389-98.

14. Kesselheim AS, Myers JA, Avorn J. Characteristics of clinical trials to support approval of orphan vs nonorphan drugs for cancer. JAMA 2011;305(22):2320-6.

15. Fojo T, Mailankody S, Lo A. Unintended consequences of expensive cancer therapeutics-the pursuit of marginal indications and a me-too mentality that stifles innovation and creativity: the John Conley Lecture. JAMA Otolaryngol Head Neck Surg 2014;140(12):1225-36.

16. Salas-Vega S, Iliopoulos O, Mossialos E. Assessment of overall survival, quality of life, and safety benefits associated with new cancer medicines. JAMA Oncol 2017;3(3):382-90.

17. Ioannidis JP. Why most discovered true associations are inflated. Epidemiology 2008;19(5):640-8.

6. Corroborating characteristics of single pivotal trial evidence supporting FDA approval of novel cancer therapies

Aviv Ladanie [1, 2], Benjamin Speich [1], Matthias Briel [1, 3], Arnav Agarwal [3, 4], John PA Ioannidis [5, 6, 7, 8, 9], Tiago V Pereira [10], Heiner C Bucher [1], Benjamin Kasenda [1, 11], Lars G Hemkens [1]

[1] Basel Institute for Clinical Epidemiology and Biostatistics, Department of Clinical Research, University Hospital and University of Basel, Basel, Switzerland
[2] Swiss Tropical and Public Health Institute (Swiss TPH), Basel, Switzerland
[3] Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Canada
[4] Department of Medicine, University of Toronto, Toronto, Ontario, Canada
[5] Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, California, USA
[6] Department of Medicine, Stanford University School of Medicine, Stanford, California, USA
[7] Department of Health Research and Policy, Stanford University School of Medicine, Stanford, California, USA
[8] Department of Biomedical Data Science, Stanford University School of Medicine, Stanford, California USA
[9] Department of Statistics, Stanford University School of Humanities and Sciences, Stanford, California, USA
[10] Health Technology Assessment Unit, Institute of Education and Health Sciences, Oswaldo Cruz German Hospital, São Paulo, Brazil
[11] Departments of Medical Oncology and of Haematology, University Hospital and University of Basel, Basel, Switzerland

Status: This manuscript is in draft form but has been approved by most co-authors.

Abstract

Introduction: Novel treatments for cancer are increasingly approved based on evidence from a single pivotal trial alone. Beyond supporting data from other studies and clinical as well as ethical implications, there are five characteristics of pivotal trial evidence used in efficacy assessments of the US Food and Drug Administration (FDA) that may indicate their higher validity and may corroborate the reliance on a single trial alone for approval decisions.

Methods: We identified all single pivotal trials supporting FDA approval of novel drugs and therapeutic biologicals for cancers between 2000 and 2016. For each trial, we determined the presence of these five characteristics, which we operationalized as (1) large and multicentric trial (200 and more patients; two or more trial centers); consistent treatment benefits across (2) multiple patient subgroups (in view of FDA reviewers), (3) multiple endpoints (including overall survival, progression-free survival, objective tumor response, and health-related quality of life) and (4) multiple treatment comparisons (e.g. multi-arm studies); (5) "statistically very persuasive" results (p-values <0.00125).

Results: Thirty-six of 100 approvals were based on evidence from a single pivotal trial (20 randomized controlled trials and 16 single-arm trials). Sixty-four percent (23/36) of single pivotal trials were large and multicentric trials (median of 316 patients and 62 centers; 23/36 (64%) had more than 200 patients, and 36/36 (100%) were multicentric). Consistent effects were demonstrated across different subgroups in 64% (23/36) of single pivotal trials, across endpoints in 42% (15/36), and across multiple comparisons in 3% (1/36). Thirty-three percent (33%, or 12/36) of single pivotal trials had very low p-values for the primary endpoint. At least one of the corroborating characteristics was present in 92% (33/36) of all single pivotal trials, two or more were present in 53% (19/36).

Conclusions and relevance: Single pivotal trials typically have some of the corroborating characteristics described by the FDA before, but half of them have only one. These characteristics may indicate a higher internal and external validity, but whether single pivotal trials with such characteristics provide similar evidence about benefits and harms of novel treatments as multiple trials would do, remains

questionable.

Keywords: US Food and Drug Administration; Drug Approval; Cancer; Confirmatory Evidence; Single Pivotal Trial; FDAMA Section 115a; Meta-Research.

Introduction

Benefits and harms of novel treatments are commonly associated with uncertainty and early evidence often provides only limited guidance to assess the true merits of such therapies (1-3). The US Congress reflected this uncertainty in the Kefauver-Harris Amendments to the Federal Food, Drug and Cosmetic Act (FFDCA) in 1962 to ensure that novel drugs are safe and effective. This established the "effectiveness requirement" for drug approvals by the US Food and Drug Administration (FDA), and stipulates "substantial evidence" from "adequate and well-controlled investigations" (4). While this has traditionally been interpreted as a general rule that at least two pivotal studies are needed for drug approval, under particular circumstances the FFDCA allows that the FDA grants approval based on early evidence from only a single pivotal trial (5). These circumstances and their legal and scientific basis are outlined in a FDA "Guidance for Industry" document (6). It summarizes several general prerequisites for such situations, for example assuming that "the single study has been appropriately designed, that the possibility of bias due to baseline imbalance, unblinding, post-hoc changes in analysis, or other factors is judged to be minimal, and that the results reflect a clear prior hypothesis documented in the protocol." It also differentiates approval situations with a single pivotal study that is accompanied by additional "supporting evidence" (from "related adequate and well-controlled studies", for example in other populations, disease stages or closely related diseases) and situations where there is only a single pivotal study existing. For the latter situation, as further prerequisite for reliance on a single trial, the FDA explicitly states that "a clinically meaningful effect on mortality, irreversible morbidity, or prevention of a disease with potentially serious outcome and confirmation of the result in a second trial would be practically or ethically impossible" has to be demonstrated.

Beyond supporting data from other studies and clinical relevance and ethical implications of the results, and possible legal circumstances, there are also some specific methodological factors of the study design and statistical results listed in the guidance document as examples of trial characteristics which may corroborate the reliance on a single pivotal trial. The corroborating characteristics specifically mentioned are (1) a large trial size and multicenteredness (with no center having disproportionately large influence on the number of enrolled patients or the overall treatment effect), treatment effects that are consistent across (2) patient subgroups, (3) multiple endpoints ("involving different events" or representing different beneficial effects), or (4) multiple comparisons (in factorial designs or multi-arm studies), and (5) "statistically very persuasive" results.

In oncology, approval based on evidence from a single pivotal study alone has become standard. Meta-epidemiological surveys show that 80 to 85% of drug approvals in oncology by the FDA and the European Medicines Agency (EMA) are based on a single pivotal study alone (Table 1). This draws attention to the abovementioned circumstances. While supporting data from other studies, clinical relevance, and ethical implications are complex and subject to numerous debates often on a case by case basis, there is, to our knowledge, no systematic evaluation of the five corroborating characteristics in pivotal oncology trials.

We aimed to focus on these situations where there is only a single pivotal study alone available at approval of novel cancer therapies and to describe their characteristics. We took for granted that all these single studies provide clinically meaningful effects and are adequate and well-controlled investigations. A better understanding of these corroborating characteristics and their prevalence may provide valuable insights in the assessment of early treatment effects and their surrounding uncertainty and also foster further meta-epidemiological analyses of this evidence.

Methods

Identification of approval evidence

We identified all 92 cancer drugs and therapeutic biologics that received a marketing authorization by the FDA for the first time (so-called new molecular entities and therapeutic biologic products, referred to as "drugs" in this article) between 2000 and 2016 in a related project (7). For each eligible drug, we obtained FDA approval packages containing reviews of the data submitted by the drug manufacturer to support their claims. These provide information about the pivotal studies and are publicly available (for approved drug products) from the drugs@FDA database (8). We described the detailed process elsewhere (9). We re-used the information about brand and generic name of the cancer drugs, FDA approved indication, approval date, orphan status, and whether a drug was granted accelerated approval from the related project. We identified eligible randomized trials in clinicaltrials.gov for cross-checks of our extractions (last search 25 July 2018).

One reviewer (AL) perused the approval packages for each of the 92 drugs. He determined the number of pivotal studies supporting the approval. This information is usually provided in either the "Executive summary" or the "Data source" chapter of medical reviews. The number of pivotal studies was classified either as "Two or more pivotal studies", "One pivotal study plus additional supporting evidence from related studies", or "Single pivotal study". A second reviewer (BK) verified these classifications in a random 30% sample. There was perfect agreement.

General characteristics of single pivotal studies

For approvals based on a single pivotal study alone, one reviewer (AL) identified the pivotal study, extracted the study name as well as the number of study arms, and categorized the study as either "Randomized controlled trial (RCT)" or "Single-arm trial (SAT)"; we use the term "trial" because all pivotal studies were experimental clinical trials. The SAT category encompasses both trials without

concurrent controls as well as dose-comparison trials where patients in all groups are receiving the experimental drug.

Corroborating characteristics of single pivotal trials

The FDA has defined five corroborating characteristics of single pivotal trials which we quote here and provide information how evidence for each of these characteristics was collected.

Large multicenter study

"*In a large multicenter study in which (1) no single study site provided an unusually large fraction of the patients and (2) no single investigator or site was disproportionately responsible for the favorable effect seen, the study's internal consistency lessens concerns about lack of generalizability of the finding or an inexplicable result attributable only to the practice of a single investigator.*" (6)

We extracted for all single pivotal trials the total number of patients enrolled, the number of participating centers, and any statement about whether a trial was dominated by a single center. The latter was usually reported in the chapter "Ethics and good clinical practices" of the medical review and is mainly derived from on-site inspection reports by the FDA's Office of Scientific Investigations and financial disclosure information from participating investigators. Since there is no established definition of what constitutes a large trial, we arbitrarily defined a trial with at least 200 enrolled patients (across all trial arms) to be large for our main analysis. The cutoff of 200 patients has been used in recent analyses by FDA scientists to distinguish between small and large cancer clinical trials (10). Trials with more than one enrolling center were considered to be multicentric. In sensitivity analyses, we used other thresholds (i.e. 500 and 1,000 patients).

Consistency across patient subgroups

"Frequently, large trials have relatively broad entry criteria and the study populations may be diverse with regard to important covariates such as concomitant or prior therapy, disease stage, age, gender or

race. Analysis of the results of such trials for consistency across key patient subsets addresses concerns about generalizability of findings to various populations in a manner that may not be possible with smaller trials or trials with more narrow entry criteria." (6)

We determined for all trials the primary endpoint used for regulatory decision-making as that one which we deemed most essential for the approval decision because it was first discussed in detail in the results section of the medical review.

Then we searched FDA documents for a discussion of subgroup effects on this endpoint, extracted the FDA's statement about the consistency across demographic and disease subgroups (for example, across age groups, sex, race, geographic region, or disease stage), and classified the statements as indicating consistency or not for this endpoint.

Consistency across multiple endpoints involving different events

"*In some cases, a single study will include several important, prospectively identified primary or secondary endpoints, each of which represents a beneficial, but different, effect. Where a study shows statistically persuasive evidence of an effect on more than one of such endpoints, the internal weight of evidence of the study is enhanced.*" (6)

We determined for all primary endpoints used for regulatory decision-making, as well as for all endpoints of overall survival (OS), progression-free survival (PFS) and objective tumor response (TR) whether the experimental treatment had more favorable effects compared to the control arm (i.e., statistically significant effects or effects crossing a pre-specified efficacy threshold). Since this information is usually only available for comparative treatment effects in FDA approval documents, we specifically assessed this characteristic only for comparative trials, i.e. RCTs (not SATs), while we deemed the other trials as not corroborated by this characteristic per se.

We extracted statements about statistical significance, p-values, alpha-levels or thresholds (where applicable) and treatment effects. For p-values reported only as being below a cut-off (e.g. $p < 0.0001$) we used this cut-off-value, and we doubled p-values from one-sided hypothesis tests to consistently

obtain two-sided p-values (11). We classified effects as statistically significant whenever explicitly declared by the FDA, or (when there was no explicit statement) when the p-value was below an explicitly defined alpha-level or below 0.05 (in cases without a defined alpha-level) or (when there also were no p-values reported), when the confidence interval excluded the null. When only one of these endpoints (i.e., the primary, OS, PFS, TR) had statistically significant favorable effects, we perused the approval documents for any statements about statistically significant favorable effects on health-related quality of life (HRQoL), if still there was only one endpoint with statistically significant favorable effects, we deemed this characteristic not met.

In a sensitivity analysis, we considered PFS (and event-free survival, the primary endpoint in the trial used for approval of UNITUXIN®, dinutuximab) not being sufficiently different to OS because survival is an integral part of the outcome definition of PFS. In these cases, we also evaluated potential benefits on HRQoL.

Consistency across treatment comparisons

"*Properly designed factorial studies may be analyzed as a series of pairwise comparisons, representing, within a single study, separate demonstrations of activity of a drug […]*" (6)

We evaluated whether a second treatment comparison within the same trial (in the FDA guidance document described as "multiple studies in a single study") confirmed a beneficial effect on the primary endpoint. Only multi-arm or factorial design RCTs were assessed for this characteristic.

Statistically very persuasive findings

"*In a multicenter study, a very low p-value indicates that the result is highly inconsistent with the null hypothesis of no treatment effect […]*" (6)

We categorized findings on the primary endpoint used for regulatory decision-making as "very persuasive" when they favored the experimental treatment with a p-value of 0.00125 or less. A p-value of 0.00125 is frequently discussed in the context of single pivotal trials (12-15) and is statistically

equivalent to running two independent trials where each trial demonstrates a statistically significant effect in favor of the experimental treatment at the conventional 0.05 level (12-15). Since it required comparative treatment effects, we also assessed this only for comparative trials (RCTs) and deemed the other trials as not corroborated by this characteristic per se.

All extractions and categorizations were conducted by one reviewer (AL). The categorizations of the first reviewer were confirmed by another reviewer (BS) and any disagreements were resolved by a third reviewer (LGH).

Cross-check with clinicaltrials.gov

One reviewer (MB) perused the information in clinicaltrials.gov (as of 25 July 2018) and determined whether in cases where OS, PFS, TR, or HRQoL was reported in the FDA documents, these outcomes were a primary or secondary endpoint of the trial (which was the case with only very few exceptions). The endpoint that we determined as primary endpoint for regulatory decision-making was always the primary endpoint of the trial (with one exception for ALIMTA® (pemetrexed) where the entry in clinicaltrials.gov was not clear).

Data analysis

We report the results descriptively using summary statistics (minimum, median, interquartile range [IQR], and maximum). We conducted no statistical hypothesis tests.

Results

A total of 92 drugs received FDA approval between 2000 and 2016 for 100 cancer indications (Figure 1). Thirty-three (33) novel drugs were approved based on evidence from 35 single pivotal trials for 36 cancer indications (two drugs were approved for 2 or 3 indications, and two indications were approved based on evidence from the same trial). The number of approvals based on a single pivotal trial

increased over time, with 67% (24/36) of indications being approved since 2012 based on evidence from single pivotal trials (Table 2). The majority of approvals (67%; 24/36) had an orphan status and numerous underwent accelerated approvals (15/36; 42%). Fifty-six percent (56%, or 20/36) of single pivotal trials were RCTs and 44% (16/36) were SATs (Table 2).

The median number of patients enrolled across the 36 single pivotal trials was 316 (IQR 125, 533). All trials were multicentric with a minimum of 16 centers (Table 2). No trial was dominated by a single center, and 23 of 36 trials enrolled more than 200 patients (9 trials had more than 500 (25%) and 4 more than 1,000 participants (11%)). The "large multicenter study" characteristic was present for 64% of all approvals (23/36) (Table 3).

Consistent treatment benefits across multiple patient subgroups were shown for 23 of 36 trials (64%) (Table 3).

The primary endpoint used for regulatory decision-making showed statistically significant benefits in all RCTs. This was PFS in 10 (50%) and OS in 7 RCTs of 20 RCTs (35%). The other 3 RCTs showed beneficial effects on TR and PFS as co-primary endpoints (TREANDA®, bendamustine hydrochloride), medical castration rate in prostate cancer (FIRMAGON®, degarelix acetate, where efficacy was established without p-value), and event-free survival (UNITUXIN®, dinutuximab, where the pivotal trial was stopped early after the 7th interim analysis indicated a p-value that was close to the pre-specified stopping boundary but did not formally cross it) to support the approval. Treatment benefits were consistent across multiple endpoints, i.e., OS, PFS, TR (and event-free survival as different primary endpoint in one case) in 15 of the 20 RCTs. Thus, of all 36 approvals based on single pivotal trial evidence (RCT or SATs), 42% (15/36) were supported by trial evidence with this corroborating characteristic (Table 3).

Only one trial (3%; 1/36) had a multi-arm or factorial design and evaluated the efficacy of two different maintenance doses against an active treatment control. Both treatment comparisons were deemed to be effective by the FDA, thus demonstrating consistency across treatment comparisons (Table 3).

For almost all RCTs small p-values were reported for the primary endpoint used for regulatory decision-making with a median two-sided p-value of 0.0001 (IQR: 0.0001, 0.0108; n=19, for one trial we

identified no p-value). For 10 RCTs, p-values were reported as "<0.0001". Overall, the p-values of 12 RCTs were smaller than 0.00125, of 13 RCTs smaller than 0.005, and of 18 RCTs smaller than 0.05 (Table 3). Of all 36 approvals based on single pivotal trial evidence (RCT or SATs), 33% (12/36) were supported by trial evidence having this corroborating characteristic (Table 3).

Overall, 92% (33/36) of all approvals based on a single pivotal trial at least one characteristic is present (Table 3). Only one characteristic was present in 14 trials (39%), two in 5 trials (14%), three in 6 trials (17%) and four characteristics were present in 8 trials (22%). There was no trial with all characteristics, but 3 trials had none (8%). The median number of the evaluated corroborating characteristics per trial was 2 (IQR 1; 3).

Using an alternative approach to define the characteristics (i.e., 1,000 patients necessary for a large trial and considering PFS and event-free survival as not sufficiently different to OS), 75% (27/36) approvals were supported by a single pivotal trial with at least one characteristic. Using a cutoff of 500 patients, results were similar (81%, 29/36; Appendix).

Discussion

FDA approval of novel drugs for cancer indications between 2000 and 2016 was based in 36 out of 100 cases on evidence from only a single pivotal trial without any further "supporting evidence" from related trials. For half of these approvals we could identify support by trials with only one of the five corroborating characteristics described by the FDA. Most approvals were supported by trials having at least one of the corroborating characteristics, but many of them by only one alone.

This systematic evaluation shows that single pivotal trials providing the only evidence for the establishment of drug efficacy were always multicentric involving often more than 30 centers with a median of 316 patients, but rarely more than 500 or 1,000 patients. Consistent effects across patient subgroups were present in most (64%) of the trials according to FDA reviewers. However, only about one-third of approvals was supported by trials with consistent effects across endpoints (42%) or by

statistically very persuasive findings (33%). The latter is mainly because these characteristics become relevant only for comparative pivotal trials, i.e. RCTs, which, however, had this characteristic in only 70% and 60%, respectively. Corroborating evidence resulting from consistency across multiple comparisons within a trial played no role because such factorial or multi-arm designs were very rare.

For three approvals (BELEODAQ®/belinostat, ERIVEDGE®/vismodegib, and ERWINAZE®/asparaginase Erwinia chrysanthemi) we found none of the corroborating characteristics that are specifically listed in the FDA guidance document. However, the guidance authors clearly emphasize that it is not "a complete listing of the circumstances in which existing efficacy data may provide independent substantiation of related claims". On the other hand, we identified no further specific criteria mentioned in the perused approval documents beyond these five characteristics. Actually, there were only two approval packages where we found a clear summary referring and discussing all of these five characteristics. For ZALTRAP® (ziv-aflibercept), for example, there is an explicit and transparent summarizing statement: "The primary issue considered during the review of this application was whether the results of a single adequate and well-controlled trial were sufficient to support approval. FDA Guidance identified characteristics that can contribute to the conclusion that results from a single study can support an efficacy claim. The characteristics identified were (a) large multi center study; (b) consistency across study subsets; (c) multiple studies in a single study; (d) multiple endpoints involving different events; and (e) statistically very persuasive findings. Results of the VELOUR trial submitted in support of this BLA satisfied all of these characteristics except (c)".

We systematically analyzed, to our knowledge for the first time, the presence of these five characteristics in single pivotal trials submitted to the FDA in support for the approval of novel drugs for cancer indications over 17 years. The number of oncological approvals supported by single trials has been explored in several other meta-epidemiological surveys before, which estimate a prevalence of 80% to 85% (16-21). This is very similar to our findings of 81% when we include the 45 of 100 indications with supportive evidence e.g., from trials in other populations. However, none of those surveys specifically focused on cases where there is no supportive evidence.

Several limitations of this study merit closer attention. First, the approval packages are large documents and very complex and despite our efforts to standardize and double-check our extractions, we cannot rule out possible extraction errors. We cross-checked the pivotal trials used in this analysis with the database of our related project (7). This data has been extensively double-checked and is based on independent extractions of at least two reviewers. We identified for three drugs further (completed) trials, but they were not relevant for the efficacy assessment (they were explicitly described as contributing only to safety analyses or as "inadequate to support efficacy claims"). We believe that any extraction errors would have minor impact and the overall interpretation would be unaffected.

Second, the five characteristics and their description in the guidance documents are rather vague, and we could not rely on a clear definition. Therefore, several of our operationalizations remain to some extent arbitrary. There is no established cutoff to distinguish between small and large trials (22), so we used a threshold of 200 patients which was used previously to estimate the robustness of effect estimates depending on the size of oncology trials (10) and selected further thresholds in sensitivity analyses. For the "multiple endpoints involving different events" characteristic, we evaluated the primary endpoints used for regulatory decision-making, systematically perused OS, PFS, and TR as the most frequently used and probably most important outcomes for approval of cancer drugs, as well as HRQoL. When reported in the FDA documents, almost all of them were primary or secondary endpoints in the trials (according to clinicaltrials.gov). However, a beneficial effect should be demonstrated across independent endpoints and it is frequently argued in discussions about outcome surrogacy that OS, PFS, and TR correlate and the FDA often accepts for example PFS as surrogate outcome of OS. We addressed this issue in a sensitivity analysis by using an alternative approach, which did not alter the main interpretation. For the "statistically very persuasive finding" characteristic, the FDA does not provide a threshold of what they consider to constitute a "very persuasive" effect. Our cutoff of 0.00125 is often proposed as the appropriate level of statistical evidence to reach for regulatory decision-making in single pivotal trials (12-15), and FDA authors stated that they "ordinarily have said that a value in the neighborhood of 0.001 is good enough for a single trial" (23). Since p-values were often reported

imprecisely as “smaller than a cut-off” and we did not know the true, smaller, p-value, our analyses may underestimate the median p-value. Since half of the endpoints were reported as “<0.0001”, even using when we had used a stricter cut-off would not have changed our interpretation.

Third, we assumed that the characteristics based on consistent effects (across endpoints and comparisons) and on “statistically very persuasive findings” do not exist in non-comparative single-arm-trials. These studies typically explore tumor response measures and no other clinical endpoints (such as OS or PFS), and typically do not report p-values for efficacy outcomes. Our approach is in line with various statements in approval documents, for example in the medical review of ERIVEDGE®/asparaginase Erwinia chrysanthemi (“Because Study SHH4476g was a single arm non-comparative study, no statistical inferential conclusion can be drawn from the study. Whether the objective response rates […] are clinically meaningful and provide a favorable benefit-risk profile […] are deferred to the clinical review team” and “[…] without a comparative arm, no inferential conclusion of statistical strength can be drawn. Descriptive statistics are more appropriate in reporting the results for a non-comparative study”.

Fourth, we focused on selected features of pivotal studies and evaluated only a fraction of factors that are relevant for approval decisions and benefit assessments of novel treatments. Since we did not evaluate treatment effects, clinical impact of findings, validity of comparators or outcomes, other types of bias, or further relevant aspects, our findings do not provide information about the benefits and harms of these drugs which would be beyond the scope of this project.

Fifth, we have not compared the prevalence of these characteristics with a control group of other studies. Finally, our findings are only applicable to applications that were finally granted approval. We cannot make any statement about applications where single pivotal trials did not suffice for licensing and how often a second pivotal trial was initiated in response to the FDA’s rejection of such applications.

Overall, we found that most of the single pivotal studies fulfill at least one of the corroborating characteristics. They often have many centers per trial but typically have less than 500 or 1,000 patients what other investigators would consider “large” (24-28). Multicenter trials provide on average smaller

effects than single-center trials (29), but smaller trials exaggerate treatment effects compared to larger trials (22, 30-34). They may be underpowered, more prone to publication bias or have lower methodological quality, which may explain some of the exaggeration (31, 32, 34, 35). One might argue that in regulatory settings publication biases play no role, methodological quality may be better and therefore this characteristic might be of less relevance. Furthermore, the rationale for this characteristic in the guidance focuses less on sample size and more on the number of centers and on the generalizability to different care settings. Our findings may indeed indicate a better generalizability of pivotal trials to different care settings than sometimes assumed, albeit we have not assessed the characteristics of the trial populations. The characteristic of consistent effects across subgroups is problematic because detection of subgroup effects requires sufficient statistical power (36) and absence of evidence for subgroup effects does not provide proof of consistent effects. Subgroup effects in trials are often not credible and spurious (37), but they may be interpreted as indications for inconsistent effects in the framework of these characteristics. For example, sex-based subgroup differences are often discussed in approval documents, but subgroup findings from individual randomized trials are rarely corroborated in meta-analyses (38). Thus, this characteristic may falsely corroborate or falsely weaken the approval evidence. The consistency of effects across outcomes is difficult to interpret as discussed above and the corroboration of the result by consistent findings in multiple comparisons is rare.

Whether single trials with these five characteristics provide similar evidence about benefits and harms of novel treatments as multiple trials would do, remains questionable. The fact that these single trials represent the first evidence on benefits also requires further consideration, since even when the studies truly show benefits, this observed benefit may be substantially inflated due to regression-to-the mean and related effects (31, 39). It also remains to be shown whether trials with such characteristics better represent the true treatment benefit than other trials. A clearer operationalization and definition of these characteristics and more structured and transparent reporting of their use for approval decisions would be helpful so that users of approval trials can better

understand and assess the reliability of the evidence.

Conclusion

Modern cancer treatments are typically evaluated in only a single pivotal trial, and frequently without any additional supporting evidence. When novel drugs for cancer indications were approved in the last 17 years by the FDA based on positive evidence from only a single pivotal trial, these trials typically have corroborating characteristics described by the FDA before. However, half of the trials had only one such characteristic and some had none. The characteristics may indicate a higher internal and external validity, but whether single trials with such characteristics provide similar evidence about benefits and harms of novel treatments as multiple trials would do, remains questionable.

Acknowledgments

Authors' contributions

AL conceived the study with input from LGH. All authors collected and verified the data. AL and LGH conducted the analyses. All authors interpreted the results. AL wrote the first draft, and all authors made critical revisions on to the manuscript. All authors read and approved the final version of the paper.

Funding

This project was supported by the Swiss Cancer League (Grant No KLS-3587-02-2015). The Basel Institute for Clinical Epidemiology and Biostatistics is supported by Stiftung Institut für klinische Epidemiologie.

Role of the funding source

None of the funders/sponsors had a role in the design and conduct of the project and preparation, review, approval of the manuscript; and decision to submit the manuscript for publication.

Data sharing

No further data available. The study dataset is available on request by the corresponding author.

Tables and figures

FIGURE 1: SELECTION FLOW DIAGRAM OF THE DIFFERENT BODIES OF EVIDENCE

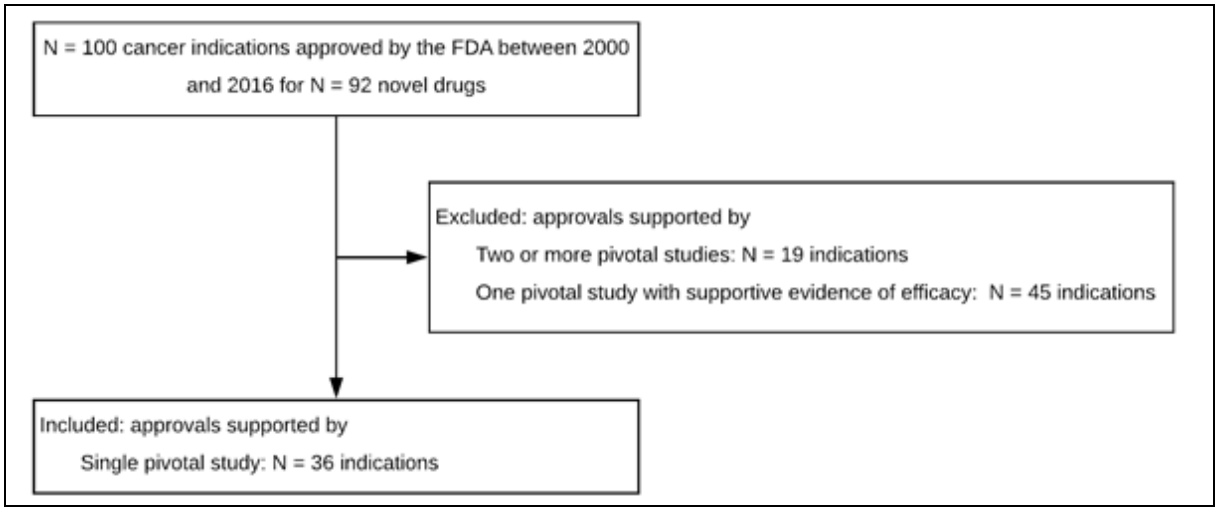


TABLE 1: STUDIES QUANTIFYING THE PREVALENCE OF SINGLE PIVOTAL STUDIES SUPPORTING DRUG APPROVALS

Study reference	Study objective(s)	Health authority analyzed	Time period analyzed *	Disease and/or drug characteristics	Prevalence of single pivotal trials supporting FDA approval, % (n/N) [analysis level]	Distinction of single trial approvals with and without supporting evidence **
Downing et al. (2014) (16)	To characterize pivotal efficacy trials for newly approved novel therapeutic agents.	FDA	2005 to 2012	Cancer (not otherwise specified)	80% (33/41) [indication]	No
Sridhara et al. (2010) (17)	To conduct an overview of products that were reviewed by the FDA's Office of Hematology and Oncology Products for marketing approval and the regulatory actions taken during July 2005 to December 2007.	FDA	2005 to 2007	Solid tumors and hematologic malignancies	83% (44/53 ***) [indication]	No
Martell et al. (2013) (18)	To describe approval trends and characteristics.	FDA	2006 to 2011	Solid tumors and hematologic malignancies	83% (NR/NR***) [indication]	No
Morant et al. (2017) (19)	To analyze the clinical efficacy evidence submitted in support of the initial marketing authorizations of new active substances approved between 2012 and 2016, with focus on approvals based on a single pivotal clinical trial.	EMA	2012 to 2016	Oncology products (according to the Anatomical Therapeutic Chemical Classification System)	84% (43/51) [drug]	No
Gentry (2015) (20)	To generate insight into the design of single pivotal studies and how the trial features were applied to provide adequate data for approval.	FDA	2005 to June 2015	Cancer (not otherwise specified, but probably both solid tumors and hematologic malignancies)	85% (55/65) [indication]	No
Tibau et al. (2017) (21)	To derive the clinically meaningful benefit for FDA approved drugs using the "European Society for Medical Oncology - Magnitude of Clinical Benefit Scale".	FDA	2006 to 2016	Solid tumors	82% (97/118 ***) [indication]	No

* Covering the time period from 1 January to 31 December if not otherwise specified; **Distinction between approvals that are genuinely based on a single trial and approvals that are based on evidence from a single pivotal trial plus supporting evidence from other studies that are unrelated to the efficacy assessment; *** including both original (i.e., the first-ever FDA-approved drug use) and supplemental approvals (i.e., changes to the original drug use approved by the FDA, such as adding a new indication. Abbreviations: EMA: European Medicines Agency; FDA: Food and Drug Administration; NR: not reported.

TABLE 2: CHARACTERISTICS OF CANCER INDICATIONS APPROVED BASED ON SINGLE PIVOTAL TRIALS

	Overall	RCTs	SATs
Indications (n, %)	36 (100)	20 (56)	16 (44)
Orphan indication (n, %)	24 (67)	10 (50)	14 (87.5)
Accelerated Approval (n, %)	15 (42)	3 (15)	12 (75)
Approval period (n, %)			
<2006 (6 years)	1 (3)	1 (5)	0 (0)
2006 - 2011 (6 years)	11 (31)	6 (30)	5 (31)
2012 - 2016 (5 years)	24 (67)	13 (65)	11 (69)
Single pivotal trial			
Median number of enrolled patients (IQR); [range]	316 (IQR 125, 533); [58 to 1226]	460 (316, 758) [133 to 1226]	123 (106, 248) [58 to 571]
Median number of participating centers (IQR); [range] *	62 (IQR 31, 114); [16 to 184]	90 (80, 147) [16 to 184]	35 (25, 50) [17 to 79]

* Information about the number of trial sites was not available for one single pivotal trial. Abbreviations: RCTs: Randomized controlled trials; SAT: Single-arm trials.

TABLE 3: CORROBORATING CHARACTERISTICS OF SINGLE PIVOTAL TRIAL EVIDENCE

Proprietary drug name (generic drug name): approved disease	Design	Large, multicenter (No of patients)	Consistency across sub-groups	Consistency across endpoints				Consistency across comparisons	Stat. very persuasive findings	Total no of characteristics
				Beneficial effects for						
				OS	PFS	RR	Multiple EP			
ADCETRIS® (*brentuximab vedotin*): anaplastic large cell lymphoma	SAT	No (58)	Yes	NA	NA	NA	NA	NA	NA	1
ADCETRIS® (*brentuximab vedotin*): hodgkin lymphoma	SAT	No (102)	Yes	NA	NA	NA	NA	NA	NA	1
AFINITOR® (*everolimus*): advanced renal cell carcinoma	RCT	Yes (416)	Yes	No	Yes[p]	No	No[2]	NA	Yes	3
ALIMTA® (*pemetrexed disodium*): malignant pleural mesothelioma, combined with cisplatin	RCT	Yes (456)	No[4]	Yes[p]	NR	Yes	Yes	NA	No	2
BELEODAQ® (*belinostat*): relapsed or refractory peripheral T-cell lymphoma	SAT	No (120)	No[4]	NA	NA	NA	NA	NA	NA	0
BOSULIF® (*bosutinib monohydrate*): relapsed or refractory chronic, accelerated or blast phase philadelphia chromosome-positive CML	SAT	Yes (571)	No[4]	NA	NA	NA	NA	NA	NA	1
COMETRIQ® (*cabozantinib s-malate*): progressive, metastatic medullary thyroid cancer	RCT	Yes (330)	Yes	No	Yes[p]	Yes	Yes	NA	Yes	4
COTELLIC® (*cobimetinib*): unresectable or metastatic melanoma with BRAF V600E or V600K mutation, combined with vemurafenib	RCT	Yes (495)	Yes	No	Yes[p]	Yes	Yes	NA	Yes	4
ERIVEDGE® (*vismodegib*): locally advanced or metastatic basal cell carcinoma	SAT	No (104)	No[4]	NA	NA	NA	NA	NA	NA	0
ERWINAZE® (*asparaginase Erwinia chrysanthemi*): acute lymphoblastic leukemia, part of multi-agent chemotherapy	SAT	No (59)	NR	NA	NA	NA	NA	NA	NA	0
FIRMAGON® (*degarelix acetate*): advanced prostate cancer	RCT	Yes (620)	NR	No	No	Yes[p]	Yes[3]	Yes	No	3
FOLOTYN® (*pralatrexate*): relapsed or refractory peripheral T-cell lymphoma	SAT	No (109)	Yes	NA	NA	NA	NA	NA	NA	1
GAZYVA® (*obinutuzumab*): CD20+ CLL, combined with chlorambucil	RCT	Yes (356)	Yes	No	Yes[p]	NR	No[2]	NA	Yes	3
GILOTRIF® (*afatinib*): EGFR mutation (exon 19 deletion or L858R)-positive metastatic non-small cell lung cancer	RCT	Yes (345)	No[4]	No	Yes[p]	NR	No[2]	NA	Yes	2
IBRANCE® (*palbociclib*): ER+/HER2- advanced breast cancer, combined with letrozole	RCT	No (165)	Yes	No	Yes[p]	NR	No[2]	NA	Yes	2
ICLUSIG® (*ponatinib* hydrochloride): chronic phase, accelerated phase, or blast phase CML	SAT	Yes (444)	NR	NA	NA	NA	NA	NA	NA	1
IMBRUVICA® (*ibrutinib*): mantle cell lymphoma	SAT	No (111)	Yes	NA	NA	NA	NA	NA	NA	1
JEVTANA KIT® (*cabazitaxel*): hormone-refractory metastatic prostate cancer	RCT	Yes (755)	Yes	Yes[p]	Yes	Yes	Yes	NA	Yes	4
KEYTRUDA® (*pembrolizumab*): unresectable or metastatic melanoma	SAT	No (173)	Yes	NA	NA	NA	NA	NA	NA	1
KYPROLIS® (*carfilzomib*): multiple myeloma	SAT	Yes (266)	NR	NA	NA	NA	NA	NA	NA	1
LARTRUVO® (*olaratumab*): soft tissue sarcoma with a histologic subtype for which an anthracycline-containing regimen is appropriate	RCT	No (133)	No[4]	Yes	Yes[p]	NR	Yes	NA	No	1
NINLARO® (*ixazomib*): multiple myeloma, combined with lenalidomide and dexamethasone	RCT	Yes (722)	No[4]	No	Yes[p]	NR	No[2]	NA	No	1
ODOMZO® (*sonidegib*): locally advanced basal cell carcinoma	SAT	Yes (230)	Yes	NA	NA	NA	NA	NA	NA	2
PORTRAZZA® (*necitumumab*): metastatic squamous non-small cell lung cancer, combined with gemcitabine and cisplatin	RCT	Yes (1093)	No[4]	Yes[p]	Yes	NR	Yes	NA	No	2
STIVARGA® (*regorafenib*): metastatic colorectal cancer	RCT	Yes (760)	Yes	Yes[p]	Yes	No	Yes	NA	No	3
TASIGNA® (*nilotinib hydrochloride monohydrate*): chronic phase and accelerated phase philadelphia chromosome-positive CML	SAT	Yes (385)	No[4]	NA	NA	NA	NA	NA	NA	1
TREANDA® (*bendamustine hydrochloride*): CLL	RCT	Yes (301)	Yes	No	Yes[p]	Yes[p]	Yes	NA	Yes	4
UNITUXIN® (*dinutuximab*): pediatric patients with high-risk neuroblastoma, combined with GM-CSF, IL-2, and 13-cis-retinoic acid	RCT	Yes (251)	Yes	No	Yes	NR	Yes[3]	NA	No	3
VECTIBIX® (*panitumumab*): EGFR-expressing, metastatic colorectal carcinoma	RCT	Yes (463)	Yes	No	Yes[p]	Yes	Yes	NA	Yes	4
VENCLEXTA® (*venetoclax*): 17p-deletion CLL	SAT	No (107)	Yes	NA	NA	NA	NA	NA	NA	1
XTANDI® (*enzalutamide*): metastatic castration-resistant prostate cancer	RCT	Yes (1199)	Yes	Yes[p]	Yes	NR	Yes	NA	Yes	4
ZALTRAP® (*ziv-aflibercept*): metastatic colorectal cancer, combined with 5-fluorouracil, leucovorin, and irinotecan	RCT	Yes (1226)	Yes	Yes[p]	Yes	Yes	Yes	NA	No	3

ZYDELIG® (*idelalisib*): relapsed CLL, combined with rituximab	RCT	Yes (220)	Yes	No	Yes[p]	Yes	Yes	NA	Yes	4
ZYDELIG® (*idelalisib*): relapsed follicular B-cell non-Hodgkin lymphoma[1]	SAT	No (125)	Yes	NA	NA	NA	NA	NA	NA	1
ZYDELIG® (*idelalisib*): relapsed small lymphocytic lymphoma[1]	SAT	No (125)	Yes	NA	NA	NA	NA	NA	NA	1
ZYTIGA® (*abiraterone acetate*): metastatic castration-resistant prostate cancer	RCT	Yes (1195)	Yes	Yes[p]	Yes	Yes	Yes	NA	Yes	4
Total n (%)		23 (64)	23 (64)	8 (22)	18 (50)	10 (28)	15 (42)	1 (3)	12 (33)	None: 3 (8%) One: 14 (39%) Two: 5 (14%) Three: 6 (17%) Four: 8 (22%) Five: 0 (0%)

1) These two cancer indications of the same novel drug are based on efficacy data from the same single pivotal trial. 2) No statements indicating benefits on quality of life found 3) characteristic met because beneficial effects also shown for (co-)primary endpoint. 4) FDA statement not conclusive or interpreted by us not to be consistent across subgroups. Abbreviations: CLL, chronic lymphocytic leukemia; CML, chronic myeloid leukemia; EP, endpoint; GM-CSF, granulocyte-macrophage colony-stimulating factor; IL, interleukin; NA, not applicable due to study design; NR, not reported; p, primary endpoint used for regulatory decision-making; RCT, randomized controlled trial; SAT, single-arm trial.

References

1. Krist AH. "Needs more research"-Implications of the Proteus effect for researchers and evidence adopters. Mayo Clin Proc 2018;93(3):273-5.
2. Naci H, Wouters OJ, Gupta R, Ioannidis JPA. Timing and characteristics of cumulative evidence available on novel therapeutic agents receiving Food and Drug Administration accelerated approval. Milbank Q 2017;95(2):261-90.
3. Mullins CD, Montgomery R, Tunis S. Uncertainty in assessing value of oncology treatments. Oncologist 2010;15 Suppl 1:58-64.
4. Kefauver-Harris Amendments of 1962. Amendment to section 505(d) of the FD&C Act (21 USC 355(d)).
5. Section 115 of the FDA Modernization Act of 1997. Amendment to section 505(d) of the FD&C Act (21 USC 355(d)).
6. US Food and Drug Administration. Guidance for industry: providing clinical evidence of effectiveness for human drug and biological products. May 1998. Available from: www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM072008.
7. Ladanie A, Speich B, Naudet F, Agarwal A, Pereira TV, Sclafani F, et al. The Comparative Effectiveness of Innovative Treatments for Cancer (CEIT-Cancer) project: rationale and design of the data collection process (manuscript accepted for publication in: Trials).
8. US Food and Drug Administration. Drugs@FDA: FDA approved drug products [Internet]. [cited 08 June 2018]. Available from: www.fda.gov/drugsatfda.
9. Ladanie A, Ewald H, Kasenda B, Hemkens LG. How to use FDA drug approval documents for evidence syntheses. BMJ 2018;362:k2815.
10. Cheng J, Zhang H, Tang S, Sridhara R. Inference based on small randomized oncology clinical trials: is the observed treatment effect true? Int J Clin Biostat Biom 2017;3:010.
11. Senn S. 2007. Statistical issues in drug development. Chapter 12.2.7: Should the two-sided p-

value always be twice the one-sided value? 2nd ed: John Wiley.

12. Deng CQ. Significant level of 0.00125. [Internet]. 23 January 2010. [cited 08 June 2018]. Available from: http://onbiostatistics.blogspot.ch/2010/01/significant-level-of-000125.html.

13. Fisher L. One large, well-designed, multicenter study as an alternative to the usual FDA paradigm. Drug Inf J 1999;33(1):265-71.

14. Senn S. 2007. Statistical issues in drug development. Chapter 12.2.8: The two-trials rule. 2nd ed: John Wiley.

15. Shun Z, Chi E, Durrleman S, Fisher L. Statistical consideration of the strategy for demonstrating clinical evidence of effectiveness-one larger vs two smaller pivotal studies. Stat Med 2005;24(11):1619-37; discussion 39-56.

16. Downing NS, Aminawung JA, Shah ND, Krumholz HM, Ross JS. Clinical trial evidence supporting FDA approval of novel therapeutic agents, 2005-2012. JAMA 2014;311(4):368-77.

17. Sridhara R, Johnson JR, Justice R, Keegan P, Chakravarty A, Pazdur R. Review of oncology and hematology drug product approvals at the US Food and Drug Administration between July 2005 and December 2007. J Natl Cancer Inst 2010;102(4):230-43.

18. Martell RE, Sermer D, Getz K, Kaitin KI. Oncology drug development and approval of systemic anticancer therapy by the U.S. Food and Drug Administration. Oncologist 2013;18(1):104-11.

19. Morant AV, Vestergaard HT. European marketing authorizations granted based on a single pivotal clinical trial: the rule or the exception? Clin Pharmacol Ther 2017;104(1):169-177.

20. Gentry L. One and done: are single pivotal studies the new norm in cancer therapeutics? [Internet]. 13 October 2015. [cited 08 June 2018]. Available from: www.ask-cato.com/one-and-done-are-single-pivotal-studies-the-new-norm-in-cancer-therapeutics/.

21. Tibau A, Molto C, Ocana A, Templeton AJ, Del Carpio LP, Del Paggio JC, et al. Magnitude of clinical benefit of cancer drugs approved by the US Food and Drug Administration. J Natl Cancer Inst 2017;110(5):486-492.

22. Ioannidis JP, Cappelleri JC, Lau J. Issues in comparisons between meta-analyses and large trials.

JAMA 1998;279(14):1089-93.

23. Temple R. How FDA currently makes decisions on clinical studies. Clin Trials 2005;2(4):276-81; discussion 364-78.

24. Sivakumar H, Peyton PJ. Poor agreement in significant findings between meta-analyses and subsequent large randomized trials in perioperative medicine. Br J Anaesth 2016;117(4):431-41.

25. Kjaergard LL, Villumsen J, Gluud C. Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. Ann Intern Med 2001;135(11):982-9.

26. LeLorier J, Gregoire G, Benhaddad A, Lapierre J, Derderian F. Discrepancies between meta-analyses and subsequent large randomized, controlled trials. N Engl J Med 1997;337(8):536-42.

27. Jones CW, Handler L, Crowell KE, Keil LG, Weaver MA, Platts-Mills TF. Non-publication of large randomized clinical trials: cross sectional analysis. BMJ 2013;347:f6104.

28. Myles PS. Why we need large randomized studies in anaesthesia. Br J Anaesth 1999;83(6):833-4.

29. Dechartres A, Boutron I, Trinquart L, Charles P, Ravaud P. Single-center trials show larger treatment effects than multicenter trials: evidence from a meta-epidemiologic study. Ann Intern Med 2011;155(1):39-51.

30. Dechartres A, Trinquart L, Boutron I, Ravaud P. Influence of trial sample size on treatment effect estimates: meta-epidemiological study. BMJ 2013;346:f2304.

31. Ioannidis JP. Why most discovered true associations are inflated. Epidemiology 2008;19(5):640-8.

32. Pereira TV, Horwitz RI, Ioannidis JP. Empirical evaluation of very large treatment effects of medical interventions. JAMA 2012;308(16):1676-84.

33. Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ 2011;343:d4002.

34. Turner RM, Bird SM, Higgins JP. The impact of study size on meta-analyses: examination of

underpowered studies in Cochrane reviews. PLoS One 2013;8(3):e59202.

35. Gluud LL, Thorlund K, Gluud C, Woods L, Harris R, Sterne JA. Correction: reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. Ann Intern Med 2008;149(3):219.

36. Brookes ST, Whitely E, Egger M, Smith GD, Mulheran PA, Peters TJ. Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test. J Clin Epidemiol 2004;57(3):229-36.

37. Sun X, Briel M, Busse JW, You JJ, Akl EA, Mejza F, et al. Credibility of claims of subgroup effects in randomised controlled trials: systematic review BMJ 2012;344:e1553.

38. Wallach JD, Sullivan PG, Trepanowski JF, Steyerberg EW, Ioannidis JP. Sex based subgroup differences in randomized controlled trials: empirical evidence from Cochrane meta-analyses. BMJ 2016;355:i5826.

39. Alahdab F, Farah W, Almasri J, Barrionuevo P, Zaiem F, Benkhadra R, et al. Treatment effect in earlier trials of patients with chronic medical conditions: a meta-epidemiologic study Mayo Clin Proc. 2018;93(3):278-83.

Appendix

SENSITIVITY ANALYSIS 1: THRESHOLD FOR "LARGE STUDY" 1,000 PATIENTS AND OS AND PFS NOT CONSIDERED INDEPENDENT

Proprietary drug name (generic drug name): approved disease	Design	Large, multicenter (No of patients) *	Consistency across sub-groups	Consistency across endpoints*				Consistency across comparisons	Stat. very persuasive findings	Total no of characteristics
				Beneficial effects for						
				OS	PFS	RR	Multiple EP			
ADCETRIS® (*brentuximab vedotin*): anaplastic large cell lymphoma	SAT	No (58)	Yes	NA	NA	NA	NA	NA	NA	1
ADCETRIS® (*brentuximab vedotin*): hodgkin lymphoma	SAT	No (102)	Yes	NA	NA	NA	NA	NA	NA	1
AFINITOR® (*everolimus*): advanced renal cell carcinoma	RCT	No (416)	Yes	No	Yes[p]	No	No[2]	NA	Yes	2
ALIMTA® (*pemetrexed disodium*): malignant pleural mesothelioma, combined with cisplatin	RCT	No (456)	No[4]	Yes[p]	NR	Yes	Yes	NA	No	1
BELEODAQ® (*belinostat*): relapsed or refractory peripheral T-cell lymphoma	SAT	No (120)	No[4]	NA	NA	NA	NA	NA	NA	0
BOSULIF® (*bosutinib monohydrate*): relapsed or refractory chronic, accelerated or blast phase philadelphia chromosome-positive CML	SAT	No (571)	No[4]	NA	NA	NA	NA	NA	NA	0
COMETRIQ® (*cabozantinib s-malate*): progressive, metastatic medullary thyroid cancer	RCT	No (330)	Yes	No	Yes[p]	Yes	Yes	NA	Yes	3
COTELLIC® (*cobimetinib*): unresectable or metastatic melanoma with BRAF V600E or V600K mutation, combined with vemurafenib	RCT	No (495)	Yes	No	Yes[p]	Yes	Yes	NA	Yes	3
ERIVEDGE® (*vismodegib*): locally advanced or metastatic basal cell carcinoma	SAT	No (104)	No[4]	NA	NA	NA	NA	NA	NA	0
ERWINAZE® (*asparaginase Erwinia chrysanthemi*): acute lymphoblastic leukemia, part of multi-agent chemotherapy	SAT	No (59)	NR	NA	NA	NA	NA	NA	NA	0
FIRMAGON® (*degarelix acetate*): advanced prostate cancer	RCT	No (620)	NR	No	No	Yes[p]	Yes[3]	Yes	No	2
FOLOTYN® (*pralatrexate*): relapsed or refractory peripheral T-cell lymphoma	SAT	No (109)	Yes	NA	NA	NA	NA	NA	NA	1
GAZYVA® (*obinutuzumab*): CD20+ CLL, combined with chlorambucil	RCT	No (356)	Yes	No	Yes[p]	NR	No[2]	NA	Yes	2
GILOTRIF® (*afatinib*): EGFR mutation (exon 19 deletion or L858R)-positive metastatic non-small cell lung cancer	RCT	No (345)	No[4]	No	Yes[p]	NR	No[2]	NA	Yes	1
IBRANCE® (*palbociclib*): ER+/HER2- advanced breast cancer, combined with letrozole	RCT	No (165)	Yes	No	Yes[p]	NR	No[2]	NA	Yes	2
ICLUSIG® (*ponatinib* hydrochloride): chronic phase, accelerated phase, or blast phase CML	SAT	No (444)	NR	NA	NA	NA	NA	NA	NA	0
IMBRUVICA® (*ibrutinib*): mantle cell lymphoma	SAT	No (111)	Yes	NA	NA	NA	NA	NA	NA	1
JEVTANA KIT® (*cabazitaxel*): hormone-refractory metastatic prostate cancer	RCT	No (755)	Yes	Yes[p]	Yes	Yes	Yes	NA	Yes	3
KEYTRUDA® (*pembrolizumab*): unresectable or metastatic melanoma	SAT	No (173)	Yes	NA	NA	NA	NA	NA	NA	1
KYPROLIS® (*carfilzomib*): multiple myeloma	SAT	No (266)	NR	NA	NA	NA	NA	NA	NA	0
LARTRUVO® (*olaratumab*): soft tissue sarcoma with a histologic subtype for which an anthracycline-containing regimen is appropriate	RCT	No (133)	No[4]	Yes	Yes[p]	NR	No[2]	NA	No	0
NINLARO® (*ixazomib*): multiple myeloma, combined with lenalidomide and dexamethasone	RCT	No (722)	No[4]	No	Yes[p]	NR	No[2]	NA	No	0
ODOMZO® (*sonidegib*): locally advanced basal cell carcinoma	SAT	No (230)	Yes	NA	NA	NA	NA	NA	NA	1
PORTRAZZA® (*necitumumab*): metastatic squamous non-small cell lung cancer, combined with gemcitabine and cisplatin	RCT	Yes (1093)	No[4]	Yes[p]	Yes	NR	No[2]	NA	No	1
STIVARGA® (*regorafenib*): metastatic colorectal cancer	RCT	No (760)	Yes	Yes[p]	Yes	No	No[2]	NA	No	1
TASIGNA® (*nilotinib hydrochloride monohydrate*): chronic phase and accelerated phase philadelphia chromosome-positive CML	SAT	No (385)	No[4]	NA	NA	NA	NA	NA	NA	0
TREANDA® (*bendamustine hydrochloride*): CLL	RCT	No (301)	Yes	No	Yes[p]	Yes[p]	Yes	NA	Yes	3
UNITUXIN® (*dinutuximab*): pediatric patients with high-risk neuroblastoma, combined with GM-CSF, IL-2, and 13-cis-retinoic acid	RCT	No (251)	Yes	No	Yes	NR	Yes[3]	NA	No	2
VECTIBIX® (*panitumumab*): EGFR-expressing, metastatic colorectal carcinoma	RCT	No (463)	Yes	No	Yes[p]	Yes	Yes	NA	Yes	3
VENCLEXTA® (*venetoclax*): 17p-deletion CLL	SAT	No (107)	Yes	NA	NA	NA	NA	NA	NA	1

XTANDI® (*enzalutamide*): metastatic castration-resistant prostate cancer	RCT	Yes (1199)	Yes	Yes[p]	Yes	NR	No[2]	NA	Yes	3
ZALTRAP® (*ziv-aflibercept*): metastatic colorectal cancer, combined with 5-fluorouracil, leucovorin, and irinotecan	RCT	Yes (1226)	Yes	Yes[p]	Yes	Yes	Yes	NA	No	3
ZYDELIG® (*idelalisib*): relapsed CLL, combined with rituximab	RCT	No (220)	Yes	No	Yes[p]	Yes	Yes	NA	Yes	3
ZYDELIG® (*idelalisib*): relapsed follicular B-cell non-Hodgkin lymphoma[1]	SAT	No (125)	Yes	NA	NA	NA	NA	NA	NA	1
ZYDELIG® (*idelalisib*): relapsed small lymphocytic lymphoma[1]	SAT	No (125)	Yes	NA	NA	NA	NA	NA	NA	1
ZYTIGA® (*abiraterone acetate*): metastatic castration-resistant prostate cancer	RCT	Yes (1195)	Yes	Yes[p]	Yes	Yes	Yes	NA	Yes	4
Total n (%)		4 (11)	23 (64)	8 (22)	18 (50)	10 (28)	11 (31)	1 (3)	12 (33)	None: 9 (25%) One: 13 (36%) Two: 5 (14%) Three: 8 (22%) Four: 1 (3%) Five: 0 (0%)

* Alternatively operationalized criterion: 1) These two cancer indications of the same novel drug are based on efficacy data from the same single pivotal trial. 2) No statements indicating benefits on quality of life found 3) characteristic met because beneficial effects also shown for (co-)primary endpoint. 4) FDA statement not conclusive or interpreted by us not to be consistent across subgroups. Abbreviations: CLL, chronic lymphocytic leukemia; CML, chronic myeloid leukemia; EP, endpoint; GM-CSF, granulocyte-macrophage colony-stimulating factor; IL, interleukin; NA, not applicable due to study design; NR, not reported; p, primary endpoint used for regulatory decision-making; RCT, randomized controlled trial; SAT, single-arm trial.

SENSITIVITY ANALYSIS 2: THRESHOLD FOR "LARGE STUDY" 500 PATIENTS AND OS AND PFS NOT CONSIDERED INDEPENDENT

				Consistency across endpoints*						
				Beneficial effects for						
Proprietary drug name (generic drug name): approved disease	Design	Large, multicenter (No of patients) *	Consistency across sub-groups	OS	PFS	RR	Multiple EP	Consistency across comparisons	Stat. very persuasive findings	Total no of characteristics
ADCETRIS® (*brentuximab vedotin*): anaplastic large cell lymphoma	SAT	No (58)	Yes	NA	NA	NA	NA	NA	NA	1
ADCETRIS® (*brentuximab vedotin*): hodgkin lymphoma	SAT	No (102)	Yes	NA	NA	NA	NA	NA	NA	1
AFINITOR® (*everolimus*): advanced renal cell carcinoma	RCT	No (416)	Yes	No	Yes[p]	No	No[2]	NA	Yes	2
ALIMTA® (*pemetrexed disodium*): malignant pleural mesothelioma, combined with cisplatin	RCT	No (456)	No[4]	Yes[p]	NR	Yes	Yes	NA	No	1
BELEODAQ® (*belinostat*): relapsed or refractory peripheral T-cell lymphoma	SAT	No (120)	No[4]	NA	NA	NA	NA	NA	NA	0
BOSULIF® (*bosutinib monohydrate*): relapsed or refractory chronic, accelerated or blast phase philadelphia chromosome-positive CML	SAT	Yes (571)	No[4]	NA	NA	NA	NA	NA	NA	1
COMETRIQ® (*cabozantinib s-malate*): progressive, metastatic medullary thyroid cancer	RCT	No (330)	Yes	No	Yes[p]	Yes	Yes	NA	Yes	3
COTELLIC® (*cobimetinib*): unresectable or metastatic melanoma with BRAF V600E or V600K mutation, combined with vemurafenib	RCT	No (495)	Yes	No	Yes[p]	Yes	Yes	NA	Yes	3
ERIVEDGE® (*vismodegib*): locally advanced or metastatic basal cell carcinoma	SAT	No (104)	No[4]	NA	NA	NA	NA	NA	NA	0
ERWINAZE® (*asparaginase Erwinia chrysanthemi*): acute lymphoblastic leukemia, part of multi-agent chemotherapy	SAT	No (59)	NR	NA	NA	NA	NA	NA	NA	0
FIRMAGON® (*degarelix acetate*): advanced prostate cancer	RCT	Yes (620)	NR	No	No	Yes[p]	Yes[3]	Yes	No	3
FOLOTYN® (*pralatrexate*): relapsed or refractory peripheral T-cell lymphoma	SAT	No (109)	Yes	NA	NA	NA	NA	NA	NA	1

GAZYVA® (*obinutuzumab*): CD20+ CLL, combined with chlorambucil	RCT	No (356)	Yes	No	Yes[p]	NR	No[2]	NA	Yes	2
GILOTRIF® (*afatinib*): EGFR mutation (exon 19 deletion or L858R)-positive metastatic non-small cell lung cancer	RCT	No (345)	No[4]	No	Yes[p]	NR	No[2]	NA	Yes	1
IBRANCE® (*palbociclib*): ER+/HER2- advanced breast cancer, combined with letrozole	RCT	No (165)	Yes	No	Yes[p]	NR	No[2]	NA	Yes	2
ICLUSIG® (*ponatinib* hydrochloride): chronic phase, accelerated phase, or blast phase CML	SAT	No (444)	NR	NA	NA	NA	NA	NA	NA	0
IMBRUVICA® (*ibrutinib*): mantle cell lymphoma	SAT	No (111)	Yes	NA	NA	NA	NA	NA	NA	1
JEVTANA KIT® (*cabazitaxel*): hormone-refractory metastatic prostate cancer	RCT	Yes (755)	Yes	Yes[p]	Yes	Yes	Yes	NA	Yes	4
KEYTRUDA® (*pembrolizumab*): unresectable or metastatic melanoma	SAT	No (173)	Yes	NA	NA	NA	NA	NA	NA	1
KYPROLIS® (*carfilzomib*): multiple myeloma	SAT	No (266)	NR	NA	NA	NA	NA	NA	NA	0
LARTRUVO® (*olaratumab*): soft tissue sarcoma with a histologic subtype for which an anthracycline-containing regimen is appropriate	RCT	No (133)	No[4]	Yes	Yes[p]	NR	No[2]	NA	No	0
NINLARO® (*ixazomib*): multiple myeloma, combined with lenalidomide and dexamethasone	RCT	Yes (722)	No[4]	No	Yes[p]	NR	No[2]	NA	No	1
ODOMZO® (*sonidegib*): locally advanced basal cell carcinoma	SAT	No (230)	Yes	NA	NA	NA	NA	NA	NA	1
PORTRAZZA® (*necitumumab*): metastatic squamous non-small cell lung cancer, combined with gemcitabine and cisplatin	RCT	Yes (1093)	No[4]	Yes[p]	Yes	NR	No[2]	NA	No	1
STIVARGA® (*regorafenib*): metastatic colorectal cancer	RCT	Yes (760)	Yes	Yes[p]	Yes	No	No[2]	NA	No	2
TASIGNA® (*nilotinib hydrochloride monohydrate*): chronic phase and accelerated phase philadelphia chromosome-positive CML	SAT	No (385)	No[4]	NA	NA	NA	NA	NA	NA	0
TREANDA® (*bendamustine hydrochloride*): CLL	RCT	No (301)	Yes	No	Yes[p]	Yes[p]	Yes	NA	Yes	3
UNITUXIN® (*dinutuximab*): pediatric patients with high-risk neuroblastoma, combined with GM-CSF, IL-2, and 13-cis-retinoic acid	RCT	No (251)	Yes	No	Yes	NR	Yes[3]	NA	No	2
VECTIBIX® (*panitumumab*): EGFR-expressing, metastatic colorectal carcinoma	RCT	No (463)	Yes	No	Yes[p]	Yes	Yes	NA	Yes	3
VENCLEXTA® (*venetoclax*): 17p-deletion CLL	SAT	No (107)	Yes	NA	NA	NA	NA	NA	NA	1
XTANDI® (*enzalutamide*): metastatic castration-resistant prostate cancer	RCT	Yes (1199)	Yes	Yes[p]	Yes	NR	No[2]	NA	Yes	3
ZALTRAP® (*ziv-aflibercept*): metastatic colorectal cancer, combined with 5-fluorouracil, leucovorin, and irinotecan	RCT	Yes (1226)	Yes	Yes[p]	Yes	Yes	Yes	NA	No	3
ZYDELIG® (*idelalisib*): relapsed CLL, combined with rituximab	RCT	No (220)	Yes	No	Yes[p]	Yes	Yes	NA	Yes	3
ZYDELIG® (*idelalisib*): relapsed follicular B-cell non-Hodgkin lymphoma[1]	SAT	No (125)	Yes	NA	NA	NA	NA	NA	NA	1
ZYDELIG® (*idelalisib*): relapsed small lymphocytic lymphoma[1]	SAT	No (125)	Yes	NA	NA	NA	NA	NA	NA	1
ZYTIGA® (*abiraterone acetate*): metastatic castration-resistant prostate cancer	RCT	Yes (1195)	Yes	Yes[p]	Yes	Yes	Yes	NA	Yes	4
Total n (%)		9 (25)	23 (64)	8 (22)	18 (50)	10 (28)	11 (31)	1 (3)	12 (33)	None: 7 (19%) One: 14 (39%) Two: 5 (14%) Three: 8 (22%) Four: 2 (6%) Five: 0 (0%)

* Alternatively operationalized criterion. 1) These two cancer indications of the same novel drug are based on efficacy data from the same single pivotal trial. 2) No statements indicating benefits on quality of life found 3) characteristic met because beneficial effects also shown for (co-)primary endpoint. 4) FDA statement not conclusive or interpreted by us not to be consistent across subgroups. Abbreviations: CLL, chronic lymphocytic leukemia; CML, chronic myeloid leukemia; EP, endpoint; GM-CSF, granulocyte-macrophage colony-stimulating factor; IL, interleukin; NA, not applicable due to study design; NR, not reported; p, primary endpoint used for regulatory decision-making; RCT, randomized controlled trial; SAT, single-arm trial.

7. General discussion

This thesis shows that pivotal clinical trials of novel cancer treatments frequently have design characteristics that may lower the confidence in their findings. Our overall results across 127 clinical trials from 2000 to 2016 are broadly comparable with the largest analysis of FDA approval evidence across various medical fields by Downing et al., which includes 55 clinical trials that supported FDA cancer drug approval between the years 2005 and 2012 (29). They found the percentage of RCTs and double-blinded trials as well as the median number of patients enrolled per trial to be lower in the therapeutic area of cancer (RCTs: 47%; double-blinding: 27%; 266 patients [IQR]: 84, 610]) compared to all therapeutic areas combined (RCTs: 89%; double-blinding: 80%; 446 patients [IQR: 205, 678]).

However, in the CEIT-Cancer dataset, the percentage of RCTs and the number of enrolled patients per trial changes notably if limited to treatments that were not approved for rare cancers or where the traditional approval was used. For these treatments, the prevalence of trials that randomly assign patients to treatment groups is similar to Downing et al.'s estimate across all therapeutic areas together (non-orphan indications: 86%; traditional approval of treatments: 84%; compared to Downing et al.'s percentage of 89% across all therapeutic areas). The same applies to the median trial size (non-orphan indications: 435 patients [IQR: 230, 760]; traditionally approved indications: 374 [IQR: 159, 710]; compared to Downing et al.'s finding of 446 patients [IQR: 205, 678] across all therapeutic areas combined). In contrast, the percentage of double-blinded trials is even lower in these strata compared to Downing et al.'s estimate (non-orphan indications: 16%; traditional approval of treatments: 27%; compared to Downing et al.'s 80% across all therapeutic areas).

Despite these differences in clinical trial characteristics, the pooled treatment effects of pivotal RCTs in the CEIT-Cancer project were broadly consistent across strata. Across all 54 pivotal RCTs, the HR for OS was 0.77 (95% CI: 0.73, 0.81) with a median survival gain of 2.40 months [IQR: 1.25, 3.89]).

We estimated the percentage of FDA drug approvals for cancer indications that are based on a single trial that provides efficacy data alone to be 36%. The number is considerably lower than the 80% to

85% estimated by others (16-21), but we only focused on studies that had no "supporting efficacy evidence" from other (related but non-pivotal) trials. A vast majority of the single pivotal trials fulfilled at least one of corroborating trial characteristics (92%) defined by the FDA, which may increase the confidence in the trial findings. In conclusion, the large percentage of indications that benefited from a more "flexible" application of the regulatory standards for drug approval is a dominant feature of novel cancer drugs and may explain the high proportion of clinical trials with less robust design characteristics in the field of cancer.

7.1. FDA drug approval standards allow a broad range of trial designs

It seems evident that the FDA's definition of "adequate and well-controlled investigations" (9) is sufficiently broad to encompass trials with a wide range of methodological design features. For example, all clinical trials with concurrent control assigned patients at random to treatment groups. However, only 51% of the trials had an internal control group, meaning that the remaining 49% were historically controlled and therefore randomization was impossible by design. This finding is surprising on first sight as legal standards for FDA drug approval require that "the method of assigning patients to treatment and control groups minimizes bias and is intended to assure comparability of the groups […] Ordinarily in a concurrently controlled study, assignment is by randomization" (9). The language of the requirement implies that randomization is not expected if a trial has no internal control, such as for historically controlled trials. Based on the data collected in the CEIT-Cancer project, it remains unanswered what measures were taken to ensure comparability between patients enrolled in the 49% single-arm trials and their historical controls and whether they were adequate. In our experience, the choice and properties of historical controls are rarely discussed in sufficient detail in medical review documents. The CEIT-Cancer database laid the foundation to closer evaluate this issue in future meta-epidemiological research.

7.2. Validity of findings is compromised or based on unverifiable assumptions

Only 51% of trials were randomized controlled, and only 24% were double-blinded. This moderate to low proportion of pivotal trials with these design characteristics is incomprehensible given the empirical evidence and theoretical arguments that support their use. Double-blinding and randomization are two key methodological principles in clinical trials that provide safeguards against confounding, performance, and detection biases (40). The Cochrane Collaboration defines performance bias as “systematic differences between groups in the care that is provided, or in exposure to factors other than the interventions of interest", and detection bias as "systematic differences between groups in how outcomes are determined" (41). Blinding of patients, investigators, and outcome assessors to the assigned treatments can minimize the impact of performance and detection bias on the overall trial findings. Properly conducted and analyzed drug efficacy trials with these two design elements are seen by many to be of the highest methodological quality (42).

A large meta-epidemiological study provides empirical evidence that RCTs without double-blinding (or in which it is unclear whether double-blinding was used) show larger treatment effects on average by 13% (ratio of odds ratio [ROR] 0.87, 95% credible interval [CrI]: 0.79, 0.96) compared to double-blind RCTs (43). It is also important to mention that the size of the bias varies with the level of subjectivity that is involved in adjudicating endpoints: mortality and other objectively assessed endpoints that require little to no interpretation by outcome assessors (such as pregnancies or laboratory endpoints) and that are therefore less prone to subjective judgments show larger treatment effects on average by only 7% to 8% (mortality, ROR = 0.92, 95% CrI: 0.80, 1.04; other objective endpoints, ROR = 0.93, 95% CrI: 0.74, 1.18). Subjectively assessed endpoints (such as pain or mental health outcomes), however, show considerably larger treatment effects by an average of 22% (ROR = 0.78, 95% CrI: 0.65, 0.92).

It is important to note that the low percentage of double-blinding of pivotal cancer trials may be less problematic if a blinded independent endpoint assessment committee is involved (44). The main purpose of these committees is to assess endpoints that may be prone to subjective judgment, such as objective tumor response to treatments (using for example radiographs to determine whether an event

occurred). These committees are not involved in the local trial conduct and have therefore no direct information about the assigned treatment, which limits the occurrence of detection bias. Whether central review committees were involved in open-label trials was not retrieved in the CEIT-Cancer project and remains unanswered.

Randomization in clinical trials is the process of randomly assigning patients to two or more treatment groups (45). It provides the best method to ensure that treatment groups are similar regarding factors that may influence the trial outcome (e.g., patient or disease characteristics such as age or stage of disease) to prevent confounding bias (46). The primary advantage of randomly allocating patients to treatment groups is that it provides reasonable assurance that unmeasured or unknown factors with a considerable influence on the trial outcome are equally distributed at baseline. In contrast, there is no such mechanism in non-randomized trials (e.g., in single-arm trials and their historical controls). Consequently, systematic differences between treatment groups in unknown factors with a considerable influence on the trial outcome cannot be verified or accounted for at the analysis stage (46). The imbalance may result in confounding that will partially or fully explain the observed difference in treatment effects. For this reason, utmost caution is warranted when interpreting findings from non-randomized trials because its validity is based on the unverifiable assumption that all important factors (that influence the trial outcome) are known and accurately considered in the analyses.

In conclusion, empirical research and theoretical considerations provide good arguments why FDA approval as well as treatment decisions of oncologists and patients should be based on evidence from clinical trials that are of the highest methodological quality.

7.3. Limited clinical relevance of treatment effects

There is no universal threshold value existing that could be used to determine whether cancer treatments achieve meaningful survival gains. Nevertheless, professional societies aimed to provide an answer to this question. The American Society of Clinical Oncology (ASCO) established four cancer-specific working groups (metastatic pancreatic, non–small-cell lung, triple-negative breast, and

colorectal cancers) consisting of patient advocates, biostatisticians, FDA and industry oncologists (47, 48). Depending on the cancer type, the four working groups considered OS HRs between 0.6 and 0.8 and median survival gains between 2.5 to 6 months to represent the minimum improvement over standard therapy needed to be considered clinically meaningful.

The European Society for Medical Oncology (ESMO) developed the "Magnitude of Clinical Benefit Scale" to determine the relative value of the various drugs indicated for the treatment of specific cancers (49, 50). It considers improvements in drug efficacy (preferably on OS), quality of life, and drug toxicity. Using information about both the relative (95% lower confidence bound of the treatment effect estimate) and absolute (gains in months) measures, the clinical value of the drug's efficacy is graded on a scale from 1 to 4 (on OS, where 1 indicates low and 4 high level of clinical value, but only those graded as 4 are considered to provide substantial improvements; for the purpose of brevity and because it is not a central element of this thesis, the detailed algorithm to determine the grade is not reproduced here).

Following the ASCO definition of clinically meaningful effects, the pooled estimate across all 54 RCTs supporting approval of various cancer types in the CEIT-Cancer project can be interpreted as borderline clinically meaningful at best (OS HR = 0.77, 95% CI: 0.73, 0.81 with a median survival gain of 2.40 months [IQR: 1.25, 3.89]). And according to the ESMO tool, the overall treatment effect achieved in the 54 RCTs would be graded 1 or 2. Overall, these results indicate that novel cancer drugs approved by the FDA between 2000 and 2016 based on evidence from RCTs provide moderate, limited, or even marginal survival gains to patients.

7.4. More transparency in regulatory decision-making needed

Always requesting evidence from trials of the highest methodological quality would bet too rigid considering the wide spectrum of diseases for which drug treatments are developed and consequently the different patient needs (18). The case of prostate cancer may illustrate the situation. The 5-year relative survival of men diagnosed between 2008 and 2014 in the US with prostate cancer in an early

stage (compared to men without the disease) is around 100% but only 30% in the more advanced (metastatic) stage (51). The FDA's position is that the latter group of patients with little or no effective treatment options may favor access to treatments that are deemed effective (based on fragile initial findings on a surrogate endpoint) over delayed availability of treatments whose benefits and harms have been better characterized in definitive clinical trials (18). On the other hand, one could argue that patients in an earlier stage of the disease and a 5-year survival rate of 100% are not in need of immediate access to new and effective treatments. One could argue that they can legitimately wait and request for the generation of high-quality evidence to make informed decisions in the future when they have reached a more advanced stage of the disease. The accelerated approval pathway tries to strike a balance between these two different demands: it provides patients with unmet medical needs with earlier access to new and potentially effective treatments and at the same time aims to ensure the generation of more comprehensive evidence about efficacy and safety of novel drugs.

This example illustrates why a single standard for drug approval may not suit all patients' needs and may provide a better understanding of the FDA's approach to apply flexibility on a case-by-case basis. But unlike for expedited programs, the circumstances under which a more flexible application of the regulatory standards is attained have not been officially formalized.

The FDA has taken initiatives for making the drug development and review process more structured and transparent by publishing "guidance for industry" documents (clarifying the FDA's current thinking on various topics) and by publishing drug approval reviews (that provide an insight into the pivotal trial evidence and the FDA's thinking that led to drug approval). However, we made the experience that many important decisions remain undisclosed. On the other hand, we reviewed in the "Corroborating characteristics of single pivotal trial evidence supporting FDA approval of novel cancer therapies" project whether the FDA's qualifying criteria for relying on evidence from a single experiment are fulfilled in the pivotal trials. Apart from the question of what constitutes a large clinical trial and transparent decision rules for determining whether there is consistency across study subsets at the FDA's side, we were generally successful in reconstructing the FDA's basis for decision-making.

7.5. Open questions and outlook

Important questions remain open. First, we focused only on the key design features of RCTs as well as their findings, although a greater insight into the design of and findings from historically controlled single-arm trials would be helpful since they supported approval of 49% of indications between 2000 and 2016. Furthermore, their share across single pivotal trials increased tremendously over time (0 out of 5 single pivotal trials between 2000 and 2006 were single-arm trials; between the years 2012 and 2016, already 11 out of 13 were single-arm trials). Their increasing importance for drug approval decisions, particularly in the field of cancer, may not be surprising in an era where targeted therapies are matched with biomarker-selected patients that are more likely to respond to the targeted treatment (52). Such trials have the potential to show large treatment effects, which, in turn, may serve as justification for the conduct of non-randomized trials. However, it would be crucial to understand whether the important methodological limitations of historically controlled trials are sufficiently addressed by sponsors and the FDA to accept findings from single-arm trials and to use these findings to guide healthcare decisions.

Second, understanding the FDA's rationale for approving novel drugs based on clinical trials that do not meet the highest methodological standards could increase public trust in the FDA. The following hypothetical questions give an insight into some of the aspects that would be of interest: "What is the distinguishing factor between an unvalidated surrogate endpoint alone and an unvalidated surrogate endpoint that is reasonably likely to predict clinical benefit and therefore may justify accelerated approval?", "How does the rarity of the disease, disease severity, or availability of effective treatment options influence acceptance of pivotal trials that do not meet the highest methodological standards?", or "How can single trials with neither of the five corroborating characteristics support FDA approval of novel cancer drugs?".

Third, a recent analysis by the FDA of all 93 accelerated approvals granted to cancer indications over 25 years shows that three indications (3%) were withdrawn because confirmatory trials failed to verify clinical benefit (53). This may be reassuring, but a recent empirical investigation focusing on surrogate

endpoints only revealed that treatment effects in clinical trials that supported FDA approval of novel drugs between 2005 and 2012 were 50% larger compared to treatment effects measured in post-approval trials with the same research question (ROR = 1.50; 95% CI: 1.01, 2.23) (54). The same remains to be answered for novel cancer drugs and endpoints measuring clinical benefit.

Future parts of the CEIT-Cancer project will aim at answering whether treatment effects estimated in pivotal cancer clinical trials are associated with larger uncertainty about drug efficacy compared to post-marketing trials. The existence and the degree of uncertainty could be evaluated in a meta-regression analysis by investigating the source of heterogeneity across the treatment effects from pre- and post-marketing trials, as well as by exploring the influence of disease and regulatory characteristics, such as cancer type, orphan status, or use of the accelerated approval pathway. This information may be helpful to get a sense of the uncertainty that is associated with early evidence, particularly for accelerated approved treatments.

7.6. Conclusion

This thesis aimed to promote the use of the drugs@FDA database in evidence syntheses of drug interventions and to describe and explain the clinical trial evidence base that supported approval of novel cancer drugs by the FDA between the years 2000 and 2016. These aims were met: the thesis explains in detail and step-by-step how to retrieve and use clinical trial information generated by biopharmaceutical companies in the pre-marketing phase of novel cancer treatments and provides an insight into the nature of this evidence base. Forty-nine percent (49%) of the clinical trials that supported FDA approval of novel drugs for cancer indications between 2000 and 2016 are designed as single-arm trials; three-quarter have an open-label design; and more than one-third of approvals are based on data generated in only a single clinical trial. The overall treatment effect across all RCTs that supported cancer drug approval provides small survival gains over control interventions.

The guide on how to use FDA drug approval documents has the potential to strengthen the validity of evidence syntheses of drug interventions by facilitating access to a data source that was repeatedly

demonstrated to be less biased compared to other information sources.

Our data and findings regarding the clinical trials supporting recent drug approvals provide a comprehensive insight into the underlying evidence of novel cancer drugs. Although the FDA's added flexibility for approving treatments for rare cancers (orphan indications) and cancers with limited treatment options based on early evidence of efficacy (accelerated approved indications) is evident from the data and was broadly expected, the detailed scientific and regulatory justifications for accepting clinical trials that do not meet the highest methodological standards remain unclear. Nevertheless, we were generally able to reconstruct the FDA's justifications for the approval of novel drugs based on evidence from a single clinical trial. It indicates that there is some transparency available to understand drug approval decisions, but still many details remain unclear. Regarding the question whether findings from historically controlled single-arm trials can be considered valid, further research is warranted.

The work and findings in this thesis will be used to inform future CEIT-Cancer projects. The database with information about the clinical trial evidence that supported FDA approval of 92 novel drugs for 100 indications will foster future meta-epidemiological investigations, including an evaluation of the amount of uncertainty that is available at approval of novel drugs.

8. References for chapters 1, 2, and 7

1. US Food and Drug Administration. What we do [Internet]. 28 March 2018. [cited 11 August 2018]. Available from: www.fda.gov/AboutFDA/WhatWeDo/.
2. US Food and Drug Administration. Investigational New Drug (IND) application [Internet]. 05 October 2017. [cited 11 August 2018]. Available from: www.fda.gov/drugs/developmentapprovalprocess/howdrugsaredevelopedandapproved/approvalapplications/investigationalnewdrugindapplication/default.htm.
3. US Food and Drug Administration. IND application reporting: protocol amendments [Internet]. 23 December 2015. [cited 11 Aug 2018]. Available from: www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/InvestigationalNewDrugINDApplication/ucm362503.htm.
4. Code of Federal Regulations (CFR). Title 21 Section 312.22 - General principles of the IND submission.
5. US Food and Drug Administration. New Drug Application (NDA) [Internet]. 29 March 2016. [cited 11 August 2018]. Available from: www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/NewDrugApplicationNDA/default.htm.
6. Code of Federal Regulations (CFR). Title 21 Section 314.50 - Content and format of an NDA.
7. US Food and Drug Administration. Development & approval process (drugs) [Internet]. 13 June 2018. [cited 11 August 2018]. Available from: www.fda.gov/drugs/developmentapprovalprocess/default.htm.
8. Katz R. FDA: evidentiary standards for drug development and approval. NeuroRx 2004;1(3):307-16.
9. Code of Federal Regulations (CFR). Title 21 Section 314.126 - Adequate and well-controlled studies.

10. US Food and Drug Administration. Guidance for industry: providing clinical evidence of effectiveness for human drugs and biological products. May 1998. Available from: www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm072008.pdf.
11. US Food and Drug Administration. FDA facts: biomarkers and surrogate endpoints [Internet]. 21 December 2017. [cited 28 March 2018]. Available from: www.fda.gov/aboutfda/innovation/ucm512503.htm.
12. Temple R. Are surrogate markers adequate to assess cardiovascular disease drugs? JAMA 1999;282(8):790-5.
13. Fleming TR, Powers JH. Biomarkers and surrogate endpoints in clinical trials. Stat Med 2012;31(25):2973-84.
14. Before Occupy: how AIDS activists seized control of the FDA in 1988 [Internet]. The Atlantic; 06 December 2011. [cited 11 August 2018]. Available from: www.theatlantic.com/health/archive/2011/12/before-occupy-how-aids-activists-seized-control-of-the-fda-in-1988/249302/.
15. Leary WE. F.D.A. pressed to approve more AIDS drugs [Internet]. The New York Times; 11 October 1988. [cited 11 August 2018]. Available from: www.nytimes.com/1988/10/11/science/fda-pressed-to-approve-more-aids-drugs.html.
16. Federal register volume 53 page 41516 (53 FR 41516, 21 October 1988).
17. Code of Federal Regulations (CFR). Title 21 Section 312.80 - Purpose.
18. Hamburg MA. Why FDA supports a flexible approach to drug development [Internet (FDA Voice blog)]. US Food and Drug Administration; 06 February 2014. [cited 11 August 2018]. Available from: https://blogs.fda.gov/fdavoice/index.php/2014/02/why-fda-supports-a-flexible-approach-to-drug-development/.
19. US Food and Drug Administration. Guidance for industry: expedited programs for serious conditions – drugs and biologics. May 2014. Available from: www.fda.gov/downloads/

Drugs/Guidances/UCM358301.pdf.

20. Code of Federal Regulations (CFR). Title 21 Sec 314.105 - Approval of an NDA and an ANDA.
21. Code of Federal Regulations (CFR). Title 21 Part 312 - Subpart E--drugs intended to treat life-threatening and severely-debilitating illnesses.
22. Code of Federal Regulations (CFR). Title 21 Section 312.81 - Scope.
23. US Food and Drug Administration. Fast track, breakthrough therapy, accelerated approval, and priority review [Internet]. 04 January 2018. [cited 11 August 2018]. Available from: www.fda.gov/ForPatients/Approvals/Fast/ucm405399.htm.
24. US Food and Drug Administration. Guidance for industry: rare diseases: common issues in drug development. August 2015. Available from: www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM458485.pdf.
25. Kesselheim AS, Wang B, Franklin JM, Darrow JJ. Trends in utilization of FDA expedited drug development and approval programs, 1987-2014: cohort study. BMJ 2015;351:h4633.
26. Light DW, Lexchin J. Why do cancer drugs get such an easy ride? BMJ 2015;350:h2068.
27. Gotzsche PC. A totally new system is needed for drug research and development. Eur J Clin Invest 2018;48(2).
28. Fojo T, Mailankody S, Lo A. Unintended consequences of expensive cancer therapeutics-the pursuit of marginal indications and a me-too mentality that stifles innovation and creativity: the John Conley Lecture. JAMA Otolaryngol Head Neck Surg 2014;140(12):1225-36.
29. Downing NS, Aminawung JA, Shah ND, Krumholz HM, Ross JS. Clinical trial evidence supporting FDA approval of novel therapeutic agents, 2005-2012. JAMA 2014;311(4):368-77.
30. Kesselheim AS, Myers JA, Avorn J. Characteristics of clinical trials to support approval of orphan vs nonorphan drugs for cancer. JAMA 2011;305(22):2320-6.
31. Sridhara R, Johnson JR, Justice R, Keegan P, Chakravarty A, Pazdur R. Review of oncology and hematology drug product approvals at the US Food and Drug Administration between July 2005 and December 2007. J Natl Cancer Inst 2010;102(4):230-43.

32. Martell RE, Sermer D, Getz K, Kaitin KI. Oncology drug development and approval of systemic anticancer therapy by the U.S. Food and Drug Administration. Oncologist 2013;18(1):104-11.

33. Morant AV, Vestergaard HT. European marketing authorizations granted based on a single pivotal clinical trial: the rule or the exception? Clin Pharmacol Ther 2017; 104(1):169-177.

34. Gentry L. One and done: are single pivotal studies the new norm in cancer therapeutics? [Internet]. 13 October 2015. [cited 08 June 2018]. Available from: www.ask-cato.com/one-and-done-are-single-pivotal-studies-the-new-norm-in-cancer-therapeutics/.

35. Tibau A, Molto C, Ocana A, Templeton AJ, Del Carpio LP, Del Paggio JC, et al. Magnitude of clinical benefit of cancer drugs approved by the US Food and Drug Administration. J Natl Cancer Inst. 2017;110(5):486-492.

36. Turner EH. A taxpayer-funded clinical trials registry and results database. PLoS Med 2004;1(3):e60.

37. Turner EH. Correction/clarification about FDA review documents. PLoS Med 2005;2(12):e422; author reply e3.

38. US Food and Drug Administration. Drugs@FDA database [Internet]. [cited 11 August 2018]. Available from: www.fda.gov/drugsatfda.

39. Schroll J, Bero L. Regulatory agencies hold the key to improving Cochrane reviews of drugs. Cochrane Database Syst Rev 2015;4:ED000098.

40. Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovitch C, Song F, et al. Evaluating non-randomised intervention studies. Health Technol Assess 2003;7(27):iii-x, 1-173.

41. Chapter 8.4: Introduction to sources of bias in clinical trials. In: Higgins J, Green S, editors. In: *Cochrane handbook for systematic reviews of interventions*: The Cochrane Collaboration, 2011. Available from: https://training.cochrane.org/handbook.

42. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, et al. Grading quality of evidence and strength of recommendations. BMJ 2004;328(7454):1490.

43. Savovic J, Jones HE, Altman DG, Harris RJ, Juni 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.

44. Ford R, Schwartz L, Dancey J, Dodd LE, Eisenhauer EA, Gwyther S, et al. Lessons learned from independent central review. Eur J Cancer 2009;45(2):268-74.

45. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928.

46. Groenwold RH, Hak E, Hoes AW. Quantitative assessment of unobserved confounding is mandatory in nonrandomized intervention studies. J Clin Epidemiol 2009;62(1):22-8.

47. Ellis LM, Bernstein DS, Voest EE, Berlin JD, Sargent D, Cortazar P, et al. American Society of Clinical Oncology perspective: raising the bar for clinical trials by defining clinically meaningful outcomes. J Clin Oncol 2014;32(12):1277-80.

48. Schnipper LE, Davidson NE, Wollins DS, Tyne C, Blayney DW, Blum D, et al. American Society of Clinical Oncology Statement: a conceptual framework to assess the value of cancer treatment options. J Clin Oncol 2015;33(23):2563-77.

49. Cherny NI, Dafni U, Bogaerts J, Latino NJ, Pentheroudakis G, Douillard JY, et al. ESMO-Magnitude of Clinical Benefit Scale version 1.1. Ann Oncol 2017;28(10):2340-66.

50. Cherny NI, Sullivan R, Dafni U, Kerst JM, Sobrero A, Zielinski C, et al. A standardised, generic, validated approach to stratify the magnitude of clinical benefit that can be anticipated from anti-cancer therapies: the European Society for Medical Oncology Magnitude of Clinical Benefit Scale (ESMO-MCBS). Ann Oncol 2015;26(8):1547-73.

51. Noone A, Howlader N, Krapcho M, Miller D, Brest A, Yu M, et al. *SEER cancer statistics review, 1975-2015*. Bethesda, MD: National Cancer Institute. https://seer.cancer.gov/csr/1975_2015/.

52. Simon R, Blumenthal GM, Rothenberg ML, Sommer J, Roberts SA, Armstrong DK, et al. The role of nonrandomized trials in the evaluation of oncology drugs. Clin Pharmacol Ther 2015;97(5):502-7.

53. Beaver JA, Howie LJ, Pelosof L, Kim T, Liu J, Goldberg KB, et al. A 25-year experience of US Food

and Drug Administration accelerated approval of malignant hematology and oncology drugs and biologics: a review. JAMA Oncol 2018;4(6):849-56.

54. Wallach JD, Ciani O, Pease AM, Gonsalves GS, Krumholz HM, Taylor RS, et al. Comparison of treatment effect sizes from pivotal and postapproval trials of novel therapeutics approved by the FDA based on surrogate markers of disease: a meta-epidemiological study. BMC Med 2018;16(1):45.