Clinical benefit, toxicity and cost of metastatic breast cancer therapies: systematic review and meta-analysis

Title: Clinical Benefit, Toxicity and Cost of Metastatic Breast Cancer Therapies: Systematic Review and Meta-analysis

Authors:
John Silberholz (corresponding author)
Ross School of Business, University of Michigan
701 Tappan Street R5468
Ann Arbor, MI 48109
josilber@umich.edu
Phone: (301) 538-0702; Fax: (734) 936-0279

Dimitris Bertsimas
Sloan School of Management, MIT
Operations Research Center, E40-111
Massachusetts Institute of Technology,
Cambridge, MA 02139

Linda Vahdat
Memorial Sloan Kettering Cancer Center
300 East 66th Street, 8th Floor
New York, NY 10065

Abstract

Purpose: Oncologists, clinical trialists, and guideline developers need tools that enable them to efficiently review the settings and results of previous studies testing metastatic breast cancer (MBC) drug therapies.

Methods: We searched the literature to identify clinical trials testing MBC drug therapies. Key eligibility criteria included at least 90% of patients enrolled in trial having MBC, therapeutic clinical trials, and Phase II-III studies. Studies were stratified based on patients' tumor receptor statuses and prior exposure to therapy. Survival and toxicity of each drug therapy were estimated from randomized controlled trials using network meta-analysis and from all studies using meta-analysis. These results, along with estimated drug costs, are presented in a web-based visualization tool.

Results: We included 1,865 studies containing 2,676 treatment arms and 184,563 patients in the tool (www.cancertrials.info). Meta-analysis-based efficacy and toxicity estimates are available for 85 HER-2-directed therapies, 84 hormonal therapies, and 442 undirected therapies. Network meta-analysis-based estimates are available for 16 HER-2-directed therapies, 26 hormonal therapies, and 131 undirected therapies.

Conclusions: In this era of increasing choices of MBC therapeutic agents and no superior approach to choosing a treatment regimen, the ability to compare multiple therapies based on survival, toxicity and cost would enable treating physicians to optimize therapeutic choices for patients. For investigators it can point them in research directions that were previously nonobvious and for guideline designers, enable them to efficiently review the MBC clinical trial

literature and visualize how regimens compare in the key dimensions of clinical benefit, toxicity, and cost.

Keywords: Metastatic breast cancer; network meta-analysis; overall survival; dose-limiting toxicity

Introduction

The literature surrounding drug therapies for cancer is large and rapidly expanding, and oncologists have limited time to stay abreast of new clinical trial results and to perform literature reviews of older articles. Tools that provide easy access to the results of the clinical trial literature are essential for practicing oncologists. In reviewing the literature, oncologists must understand the population tested in a study, as well as the efficacy and toxicity outcomes. Further, it is increasingly necessary to integrate drug costs into the relative risks and benefits of a treatment regimen.

Treatment guidelines such as the ones published by the American Society for Clinical Oncology (ASCO) [1] and the National Comprehensive Cancer Network (NCCN) [2], enable oncologists to efficiently identify state-of-the-art treatments for different patient subpopulations, are often developed based on literature reviews and panel discussions of oncology experts, and typically reference large-scale randomized controlled trials (RCTs) to support recommendations. Guidelines do not aim to present all clinical trials of recommended or non-recommended therapies, and do not always capture the range of outcomes obtained in trials testing therapies in different dosages and schedules and in different populations.

Systematic reviews identify and present results from clinical trials that address a similar research question, often including a meta-analysis or network meta-analysis to pool the results of individual studies into a more precise combined estimate. A number of network meta-analyses have compared treatment outcomes for metastatic breast cancer drug therapies targeting different patient subpopulations, like targeted agents plus chemotherapy for advanced triple-negative breast cancer [3], therapies for first- and second-line treatment of hormone-receptor-positive, HER-2-negative advanced breast cancer [4], or combination therapies for first-line treatment of HER-2-positive metastatic breast cancer [5]. These analyses typically focus on a subset of metastatic breast cancer treatments and focus solely on randomized controlled trials.

Beyond practicing oncologists, clinical trialists and treatment guidelines designers also rely on efficiently reviewing and analyzing the clinical trial literature. Trialists rely on literature reviews to understand the outcomes of previous clinical trials testing drug combinations, which can guide their own design of drug therapies. Participants in the treatment guideline design process also rely on literature reviews across a wide range of different types of therapies, and the ASCO and NCCN guideline design processes both rely on information from systematic reviews, which can be time consuming to complete.

In this work, we perform a systematic review of Phase II-III clinical trials of drug therapies for metastatic breast cancer (MBC), identifying both randomized and non-randomized studies. We present information from these studies in a web-based visualization tool, which presents aggregate information about the clinical benefit and toxicity observed in clinical trials testing different drug therapies, as well as the costs of the drugs tested in those studies.

Additionally, the tool presents results of individual studies, capturing the outcomes of drug therapies in different populations and dosage schedules. We present case studies to illustrate how the visualization tool could benefit practicing oncologists and clinical trialists.

Methods

Systematic Literature Review and Data Extraction

We performed a systematic literature review; the review protocol was not registered. We included English-language reports of clinical trials testing drug therapies for advanced or metastatic breast cancer. To obtain a sufficiently homogeneous set of trials with a clear set of drugs and dosages being tested, exclusion criteria were:

- Phase I clinical trials, trials with fewer than 10 patients treated at any dosage level (or for which these patient counts are not reported), or dose escalation studies
- Phase IV clinical trials, compassionate use programs, or observational trials – though Phase IV studies provide invaluable real-world evidence, patient characteristics and counts are sufficiently different from Phase II-III trials that we excluded them to avoid introducing systematic differences in clinical benefit and toxicity estimates based on whether a Phase IV study had occurred
- Trials testing either non-drug therapies, drugs to treat cancer side effects (e.g. therapy for painful bone metastases), or drugs to treat therapy side effects (e.g. granulocyte colony stimulating factor)

- Trials testing sequential therapies in which patients transition from one treatment to another after a pre-specified number of treatment cycles (trials removing a drug upon reaching a pre-specified number of cycles or cumulative dosage were allowed)
- Trials reporting that less than 90 percent of patients had metastatic disease

In October 2012, we queried the Cochrane Central Register of Controlled Trials with MeSH term "Breast Neoplasms" and qualifier "drug therapy." Additionally, we searched Pubmed in October 2012 with *"breast" in title AND ("advanced" OR "metastatic") in title AND ("trial" OR "phase") in title*. In July 2013, November 2015, January 2017, June 2017, March 2018, and March 2019, we searched Pubmed with *(Breast Neoplasms/drug therapy[MAJR] OR Breast Neoplasms/drug therapy[MeSH Terms]) AND (Clinical Trial[ptyp] OR Phase[Title])*.

Titles and abstracts were reviewed for inclusion/exclusion by the study team with assistance from several other individuals,[1] when necessary accessing the article's main text. The same team extracted from each identified treatment arm the fraction of patients with: metastatic disease, postmenopausal status, and prior chemotherapy, HER-2-directed therapy, or hormonal therapy (in any setting or in the palliative setting). Additional demographic factors were average age and mean ECOG performance status (PS). Additionally, the team extracted the fraction of patients with positive steroid hormone receptor status and HER-2-positive disease. Finally, the team extracted the number of patients receiving the therapy, the drugs tested and their dosing schedules, the median overall survival, the median progression-free survival (PFS) and/or time to progression (TTP), and the proportion of patients with each Grade

[1] Assistance in the literature review and data extraction was provided by former MIT graduate students Allison O'Hair and Stephen Relyea and by MIT and Wellesley undergraduate students Emily Chen, Michael Chen, Shahrin Islam, Siva Nagarajan, David Sukhin, Pei Tao, Roza Trilesskaya, Victoria Wang, Mimi Williams, and Joanna Yeh.

3 or 4 toxicity, using the NCI CTCAE v3 scale [6]. Loading doses and dosage reductions due to toxicity were not extracted, and the most commonly used dosage was extracted from studies with a mid-trial dose modification. From Phase I/II studies, only Phase II data were used.

Clinical Benefit, Toxicity, and Cost Outcomes

Clinical benefit measures of median overall survival (OS) and median PFS/TTP (all treatment arms) and OS hazard ratio and PFS/TTP hazard ratio (RCTs) were extracted using standard techniques [7].

We measure each treatment arm's toxicity using dose-limiting toxicity (DLT). We define a patient to have experienced a DLT if they had a Grade 3 or 4 non-hematologic toxicity (excluding alopecia, nausea, and vomiting) or a Grade 4 hematologic toxicity. We estimate this value based on the proportion of patients experiencing grade 3 or 4 toxicities of different types using the technique from [8-9], as detailed in the online supplement. Trials that did not specify toxicity grades or that reported toxicity by cycle were labeled as missing the DLT proportion outcome. For randomized comparisons, we extract the risk difference of a DLT.

Among RCTs with at least two eligible arms, we assess the proportion that are double blinded and that report OS hazard ratio, PFS hazard ratio, and DLT risk difference. Unblinded RCTs are considered at increased risk of bias in subjective toxicity reporting. Among all arms, we report the proportion that report median OS, median PFS/TTP, and the DLT rate. These measures of risk of bias inform the discussion of limitations of this study but are not used in data synthesis.

We measure the monthly cost of the drugs in a therapy in the United States using the Medicare Part B Drug and Biological Average Sales Price from April 2017 for the drugs covered under the Medicare program, and otherwise using the lowest available price as of June 2017 on the website www.goodrx.com, a prescription drug pricing website. Dosing details are described in the supplemental appendix.

Statistical Methods

All statistical analyses are stratified by receptor status and prior therapy. We define 14 strata; each arm belongs to at least one. Three strata capture tumor receptor status: arms for which at least 95% of patients are HER-2 overexpressing/amplified or for which the inclusion criteria required HER-2-overexpressing/amplified patients, arms for which at least 95% of patients are hormone receptor positive, and arms for which neither is true (termed "undirected"). For these three strata, drug combinations were labeled based on whether they are recommended for treatment in that patient population in the NCCN treatment guidelines [10]. Treatment arms with HER-2-positive patients are further classified into four sub-strata: $\geq$ 90% of patients having prior HER-2-directed therapy, $\leq$10% with prior HER-2-directed therapy, $\geq$90% with prior palliative HER-2-directed therapy, and $\leq$10% with prior palliative HER-2-directed therapy. We similarly defined three additional strata for arms with hormone receptor-positive patients based on hormonal therapy (the prior palliative hormonal therapy stratum was removed due to small study count) and four additional strata for undirected arms based on chemotherapy. Imputation of hormone receptor status is described in the online supplement.

We summarize evidence from RCTs via random-effects frequentist network meta-analysis (NMA), performing a separate NMA for OS hazard ratio, PFS/TTP hazard ratio, and DLT risk difference in all 14 strata. The random effects variance was estimated using the method of moments [11], and study outcome heterogeneity was quantified with I^2 [12]. For each stratum, the NMA reference treatment was a therapy with high degree in the OS, PFS/TTP, and DLT evidence networks. Only the connected component containing the reference treatment was included in the NMA.

For each stratum, we summarize evidence from all arms testing each drug combination using random-effects meta-analysis applied to the median OS, the median PFS/TTP, and the DLT rate. The variance of each study's median OS and PFS/TTP were estimated under the assumption of exponentially distributed survival; details appear in the supplemental appendix. The random effects variance for drug combinations tested in two or more studies in a stratum was estimated using the method of DerSimonian and Laird [11]. Study outcome heterogeneity was quantified with I^2 [12].

Additionally, for each stratum we perform random-effects meta-analysis of trial results standardized to a typical study population in that stratum. For each stratum, we build linear models of the median OS, median PFS/TTP, and DLT rate in a treatment arm, with observations weighted by the number of patients in each arm. Independent variables include indicator variables for each drug tested in the trial arm and potential confounders of the relationship between the drug therapy given and the patient outcomes: median age, mean performance status, study year, and proportion of patients with: prior chemotherapy, prior hormonal therapy (HR-positive strata), prior HER-2-directed therapy (HER-2-positive strata), visceral or

metastatic disease, and premenopausal status. For prior chemotherapy, hormonal therapy, and HER-2-directed therapies, we separately capture prior palliative therapy and prior therapy in any setting. Variables are only included if at least five observations differ from the median value in the stratum. Due to high variable count, models are fitted using lasso [13], with the regularization parameter selected via 10-fold cross-validation. Generally, missing independent variable values are mean imputed; imputation details are provided in the online supplement. We use the lasso models to estimate standardized clinical benefit and toxicity outcomes for each study arm, using the coefficients for each potential confounder to estimate the study outcomes if those independent variables each took their average value for that stratum.

To assess out-of-sample performance, we obtained meta-analysis estimates of the standardized median OS, median PFS/TTP, and DLT rate for all combinations using data through the March 2018 data pull. For new study arms in the March 2019 data pull, we assessed whether all 95% confidence intervals of the random effects meta-analyses for each relevant stratum overlapped the new study arm's 95% confidence interval for the outcome.

All computations used R version 3.4.3, using packages caret (version 6.0-77), glmnet (version 2.0-13), and netmeta (version 0.9-6). The netmeta package approach has been proven equivalent to frequentist network meta-analysis [14]. The web-based visualization was implemented using javascript visualization package d3.js.

Results

The literature review is summarized in Figure 1. In total 2,676 treatment arms from 1,865 studies (68% non-randomized) containing 184,563 patients were identified. Table 1 summarizes study characteristics for each stratum, and Supplemental Table 1 provides references, strata, patient counts, and outcome measures for each study arm.

Table 1 shows that arms with HER-2-positive or hormone receptor-positive patients had higher median OS than other arms, as did arms in which patients had no previous treatment. Arms with hormone receptor-positive populations had better toxicity outcomes than others.

Forty two network meta-analyses were performed, for OS hazard ratio, PFS/TTP hazard ratio, and DLT risk difference in all 14 strata. Table 2 summarizes the number of RCTs and drug therapies analyzed in each network meta-analysis and the heterogeneity encountered, and Supplemental Table 2 provides the estimated OS hazard ratio, PFS/TTP hazard ratio, and DLT risk difference for each drug combination in each stratum compared to the reference drug therapy. The observed heterogeneity/inconsistency was moderate for OS hazard ratio (average I^2=22%), PFS/TTP hazard ratio (average I^2=25%), and toxicity risk differences (average I^2=29%); the largest heterogeneity was found in toxicity comparisons for HER-2-directed therapy (I^2=76%), toxicity comparisons for undirected therapy (I^2=79%), and several sub-strata. Estimates of clinical benefit (OS and/or PFS/TTP hazard ratio) and DLT risk difference could both be computed for 16 HER-2-directed therapies, 26 hormonal therapies, and 131 undirected therapies.

Risk of bias was assessed for 561 RCTs; Supplemental Table 3 provides detailed results. Few studies (15%) were fully double blinded when assessing the outcomes measures; hormonal therapies had the highest rate of fully double-blinded studies (43%), and rates were lower for

undirected therapies (7%) and HER-2-directed therapies (14%). In total, 66% of studies fully reported overall survival hazard ratios, 58% fully reported PFS/TTP hazard ratios, and 59% fully reported DLT risk differences; undirected therapies (60%) and HER-2-directed therapies (77%) had higher DLT reporting rates than hormonal therapies (50%).

DLT rates, median OS values, and median PFS/TTP values were aggregated in 3,196 and 2,794, and 2,764 meta-analyses, respectively, as summarized in Table 3 and Supplemental Table 4. Median OS and median PFS/TTP exhibited moderate heterogeneity, with average I^2 33% (inter-quartile range [IQR] 0-66%) and 40% (IQR 0-74%), respectively, while DLT rate exhibited moderate-to-high heterogeneity, with average I^2 65% (IQR 40-93%). Adjustment of study outcomes to standard populations yielded a 4% average decrease in I^2 for median OS meta-analyses ($p=0.002$) and no statistically significant change in the I^2 for the DLT or median PFS/TTP meta-analyses. Estimates of clinical benefit (median OS and/or median PFS/TTP) and DLT rate could both be computed for 85 HER-2-directed therapies, 84 hormonal therapies, and 442 undirected therapies. Median OS, median PFS/TTP, and DLT rate were reported in 62%, 62%, and 67% of arms, respectively; undirected therapies (66%) and HER-2-directed therapies (84%) had higher DLT reporting rates than hormonal therapies (58%).

The March 2019 data pull from Pubmed yielded 30 new studies and 45 study arms. In 15 cases, a new study arm reported a survival outcome for a drug combination that had a meta-analysis estimate based on studies obtained through the March 2018 data pull. In 13 of these arms (87%), the 95% confidence intervals for the standardized outcome from the new study arm overlapped the 95% confidence intervals from the meta-analyses for all survival outcomes in all relevant strata. The largest gaps between confidence intervals (up to 6.7 months for

median OS and 3.5 months for median PFS/TTP) occurred for the control arm of [15], which differed from most other nab-paclitaxel monotherapy arms in requiring triple negative breast cancer status; no other gap exceeded 1.4 months (median OS) or 0.3 months (median PFS/TTP). Similarly, 13 of 16 study arms (81%) had overlapping confidence intervals for the standardized DLT rate in all strata. The largest gaps between confidence intervals (up to 0.32) occurred in [16], which differed from the two other letrozole-palbociclib arms in its Japanese patient population; no other gap exceeded 0.09.

The visualization tool interactively displays summary information about therapies for different patient subpopulations. For instance, Figure 2 displays NMA estimates of OS hazard ratios and DLT risk differences of drug combinations compared to cyclophosphamide / doxorubicin / fluorouracil therapy in the undirected stratum, with color representing the monthly drug costs of the therapy. Figure 3 displays standardized estimates of the toxicity and efficacy of drug combinations for the HER-2-positive patient population, with color representing whether each therapy is recommended in the NCCN treatment guidelines. The tool drills down to details about the patient population, dosages tested, and outcomes of specific trials, as displayed in Figure 4. The website is publicly available at cancertrials.info.

Discussion

The proposed visualization tool enables users to efficiently identify and understand typical outcomes for clinical trials testing different drug therapies for MBC, as well as costs for those therapies. By summarizing both randomized and non-randomized studies, the tool

captures a range of dosages, dosing schedules, and patient populations for different therapies. We believe this tool will speed literature reviews of MBC drug therapies and help users synthesize the MBC literature.

Practicing oncologists could use the tool to summarize the literature and enhance clinical decision making. As an example, consider an oncologist selecting a first-line hormonal therapy for a hormone receptor-positive patient. There is no established optimal hormonal therapy sequence for MBC patients. Options include selective estrogen receptor modulators, aromatase inhibitors, and ER downregulators with or without a CDK4/6 inhibitor. The tool shows that for hormone receptor positive patients with no prior palliative hormonal therapy (i.e., first-line therapy for metastatic breast cancer), anastrazole and letrozole have been broadly tested and have similar adverse event profiles and overall survival outcomes when evaluating using the meta-analysis with demographic adjustment. Interestingly, the letrozole and palbociclib combination has a similar median OS but higher toxicity and cost. Meanwhile, the median PFS/TTP outcome is dramatically larger for letrozole and palbociclib compared to the others. From the study-level results, only the PALOMA-1 study has reported median OS for letrozole-palbociclib [17], so mature median OS data from the larger PALOMA-2 study [18] may impact the tool's estimate of this doublet's median OS; median PFS/TTP information is already available from PALOMA-2.

Clinical trialists could also benefit from the tool. As an example, consider a trialist interested in designing capecitabine-based regimens for elderly patients with triple negative MBC. By filtering to drug therapies containing capecitabine in the undirected stratum, the tool displays multiple broadly tested capecitabine-based therapies with similar median overall

survival and toxicity outcomes: capecitabine monotherapy and doublets of capecitabine combined with cyclophosphamide, vinorelbine, paclitaxel, bevacizumab, and docetaxel. Meanwhile, two broadly tested therapies are significantly more toxic and therefore less ideal for elderly populations: bevacizumab-capecitabine-paclitaxel and capecitabine-ixabepilone. Delving into the study results for the six less toxic regimens, these regimens have been tested in 13 recent trial arms published since 2015. Further, two studies tested capecitabine in elderly populations: a non-randomized study of capecitabine-vinorelbine [19] and a randomized study of capecitabine vs. pegylated liposomal doxorubicin [20]. Reviewing these studies and their findings would be critical for our trialist. The dosing information provided by the visualization tool identifies a number of studies with daily capecitabine dosages well below standard dosages of 1,650-2,500 mg/m^2. Studying the outcomes of these identified trials may be of particular interest to our trialist.

This work has a number of limitations. First, we limited to published clinical trial reports, which may introduce publication bias [21] and miss survival updates only published in conference abstracts. The use of Pubmed MeSH terms in the search strategy also limits the ability to find very recent articles. The low rate of double blinding in RCTs cause some concern of bias for subjective toxicity outcomes. The relatively low rates of reporting OS hazard ratios and DLT risk differences may indicate reporting bias that might favor experimental treatments. Further, care must be taken in interpreting aggregated estimates of clinical benefit and toxicity obtained by averaging unadjusted or standardized results from clinical trial arms. While the standardization procedure used in this work controls for a number of important covariates, data density was insufficient to include a number of other covariates known to impact clinical

benefit in patients, such as response to prior treatment, number of prior therapies for advanced disease, disease-free interval, and disease symptoms at baseline. This unobserved confounding can bias the resulting estimates of median overall survival, median PFS/TTP, and DLT proportion for drug combinations. Deviations from the assumptions used in median survival variance estimation may lead to incorrect weights being used in meta-analyses. Further, studies where the median overall survival was not yet reached were not included in the survival estimates, potentially introducing downward bias on median overall survival estimates; this bias could be significant given that 42% of study arms did not report median overall survival. Finally, different patients experience different drug costs based on their insurance and geography, an effect that is not captured in the reported cost information in the tool.

We plan to incorporate monthly updates to the data displayed at cancertrials.info. We believe it would be beneficial to expand the scope of this effort to create publicly accessible visualization tools of the results of clinical trials for drug therapies of other diseases, both cancer and otherwise.

Data Availability: The datasets analyzed during the current study are available online at cancertrials.info.

Funding: There were no funding sources for this work.

Compliance with Ethical Standards

Conflict of Interest: Linda Vahdat has performed consulting for Berg Pharma, Seattle Genetics, Athenex, and Eisai and has received clinical trial research funding from Genentech and

Immunomedics. Dimitris Bertsimas and John Silberholz declare that they have no conflict of interest.

Ethical Approval: This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Schnipper L, Davidson N, Wollins D et al (2016) Updating the American Society of Clinical Oncology value framework: revisions and reflections in response to comments received. J Clin Oncol 34(24):2925-2934
2. National Comprehensive Cancer Network (2016) NCCN clinical practice guidelines in oncology (NCCN Guidelines®) with NCCN Evidence Blocks™. https://www.nccn.org/EvidenceBlocks/. Accessed 28 December 2016
3. Ge L, Tang Y, Zhang Q, et al (2017) A network meta-analysis on the efficacy of targeted agents in combination with chemotherapy for treatment of advanced/metastatic triple-negative breast cancer. Oncotarget 8(35):59539-59551
4. Wilson F, Varu A, Mitra D, Cameron C, Iyer S (2017) Systematic review and network meta-analysis comparing palbociclib with chemotherapy agents for the treatment of postmenopausal women with HR-positive and HER2-negative advanced/metastatic breast cancer. Breast Cancer Res Treat 166(1):167-177
5. Leung HWC, Leung JH, Chan ALF (2018) Efficacy and safety of a combination of HER2-targeted agents as first-line treatment for metastatic HER2-positive breast cancer: a network meta-analysis. Expert Opin Drug Saf 17(1):1-7

6. Trotti A, Colevas AD, Setser A, et al (2003) CTCAE v3.0: development of a comprehensive grading system for the adverse effects of cancer treatment. Semin Radiat Oncol 13(3):176-181
7. Parmar MK, Torri V, Stewart L (1998) Extracting summary statistics to perform meta-analysis of the published literature for survival endpoints. Stat Med 17(24):2815-2834
8. Silberholz J (2015) Analytics for improved cancer screening and treatment. Dissertation, MIT
9. Bertsimas D, O'Hair A, Relyea S, Silberholz J (2016) An analytics approach to designing combination chemotherapy regimens for cancer. Management Science 62(5):1511-1531
10. National Comprehensive Cancer Network (2016) NCCN clinical practice guidelines in oncology: breast cancer, version 2.2016
11. DerSimonian R, Laird N (1986) Meta-analysis in clinical trials. Control Clin Trials 7:177-188
12. Higgins J, Thompson S, Deeks J, Altman D (2003) Measuring inconsistency in meta-analyses. BMJ 327(7414):557-560
13. Tibshirani R (1996) Regression shrinkage and selection via the lasso. J R Statist Soc B 58(1):267-288
14. Rücker G, Schwarzer G (2014) Reduce dimension or reduce weights? Comparing two approaches to multi-arm studies in network meta-analysis. Stat Med 33(25):4353-4369
15. Schmid P, Adams S, Rugo HS, et al (2018) Atezolizumab and Nab-Paclitaxel in Advanced Triple-Negative Breast Cancer. N Engl J Med 379(22): 2108-2121

16. Masuda N, Nishimura R, Takahashi M, et al (2018) Palbociclib in combination with letrozole as first-line treatment for advanced breast cancer: A Japanese phase II study. Cancer Sci 109(3): 803-813
17. Finn R, Crown J, Lang I, et al (2015) The cyclin-dependent kinase 4/6 inhibitor palbociclib in combination with letrozole versus letrozole alone as first-line treatment of oestrogen receptor-positive, HER2-negative, advanced breast cancer (PALOMA-1/TRIO-18): a randomised phase 2 study. Lancet Oncol 16(1): 25-35
18. Finn R, Martin M, Rugo H, et al (2016) Palbociclib and letrozole in advanced breast cancer. N Engl J Med 375(20): 1925-1936
19. Hess D, Köberle D, Thürlimann B, et al (2007) Capecitabine and vinorelbine as first-line treatment in elderly patients (≥65 years) with metastatic breast cancer: a phase II trial (SAKK 25/99). Oncology 73(3-4):228-237
20. Smorenburg CH, de Groot SM, van Leeuwen-Stok AE, et al (2014) A randomized phase III study comparing pegylated liposomal doxorubicin with capecitabine as first-line chemotherapy in elderly patients with metastatic breast cancer: results of the OMEGA study of the Dutch Breast Cancer Research Group BOOG. Ann Oncol 25(3):599-605
21. Easterbrook PJ, Gopalan R, Berlin JA, Matthews DR (1991) Publication bias in clinical research. Lancet 337(8746):867-872

Tables and Figures

Fig. 1 Flow chart of the literature review for included studies

Stratum	Num. Treatment Arms	Median Age	Mean ECOG PS	Prop. with visceral disease	Median OS (mos)	Median PFS / TTP (mos)	Prop. with a DLT
HER-2+	228	53.9	0.48	0.68	27.4	8.9	0.32
HER-2+ w/o palliative anti-HER-2 therapy	116	54.3	0.47	0.68	32.0	10.7	0.33
HER-2+ w/ palliative anti-HER-2 therapy	28	53.3	0.50	0.68	21.0	6.0	0.27
HER-2+ w/o anti-HER-2 therapy	61	53.6	0.53	0.69	26.7	8.9	0.34
HER-2+ w/ anti-HER-2 therapy	71	53.4	0.48	0.67	21.1	6.1	0.31
HR+	349	61.7	0.59	0.50	28.3	7.7	0.15
HR+ w/o palliative hormonal therapy	110	61.8	0.58	0.47	34.9	11.0	0.14
HR+ w/o hormonal therapy	46	61.6	0.66	0.38	31.5	9.1	0.07
HR+ w/ hormonal therapy	158	62.4	0.58	0.53	24.3	5.3	0.15
Undirected	2105	55.6	0.67	0.61	17.0	6.6	0.36
Undirected w/o palliative chemo.	774	56.4	0.64	0.60	20.3	8.2	0.40
Undirected w/ palliative chemo.	317	54.5	0.75	0.66	12.5	4.6	0.33
Undirected w/o chemo.	173	57.0	0.74	0.52	16.9	7.4	0.24
Undirected w chemo.	855	54.0	0.71	0.65	13.4	5.1	0.35

Table 1: Average patient demographic values of treatment arms in the 14 strata, weighted by the number of patients in each arm.

Stratum	Dose-Limiting Toxicity		Median PFS/TTP		Median Overall Survival	
	Num. Combo	I^2	Num. Combo	I^2	Num. Combo	I^2
HER-2+	17	76	20	61	25	36
HER-2+ w/o palliative anti-HER-2 therapy	7	90	11	52	11	14
HER-2+ w/ palliative anti-HER-2 therapy	4	0	4	0	4	0
HER-2+ w/o anti-HER-2 therapy	3	0	5	0	5	0
HER-2+ w/ anti-HER-2 therapy	6	0	6	0	6	47
HR+	26	16	48	57	42	27
HR+ w/o palliative hormonal therapy	10	0	25	45	24	0
HR+ w/o hormonal therapy	3	0	10	0	14	0
HR+ w/ hormonal therapy	17	0	25	0	19	0
Undirected	147	79	119	43	165	32
Undirected w/o palliative chemo.	94	71	74	41	108	35
Undirected w/ palliative chemo.	7	0	6	0	5	0
Undirected w/o chemo.	11	0	3	0	32	65
Undirected w/ chemo.	57	79	49	46	52	45

Table 2: Summary of network meta-analyses performed in each of the 14 strata. The number of drug combinations compared in each NMA are reported in the Num. Combo column, along with the I^2 of each NMA.

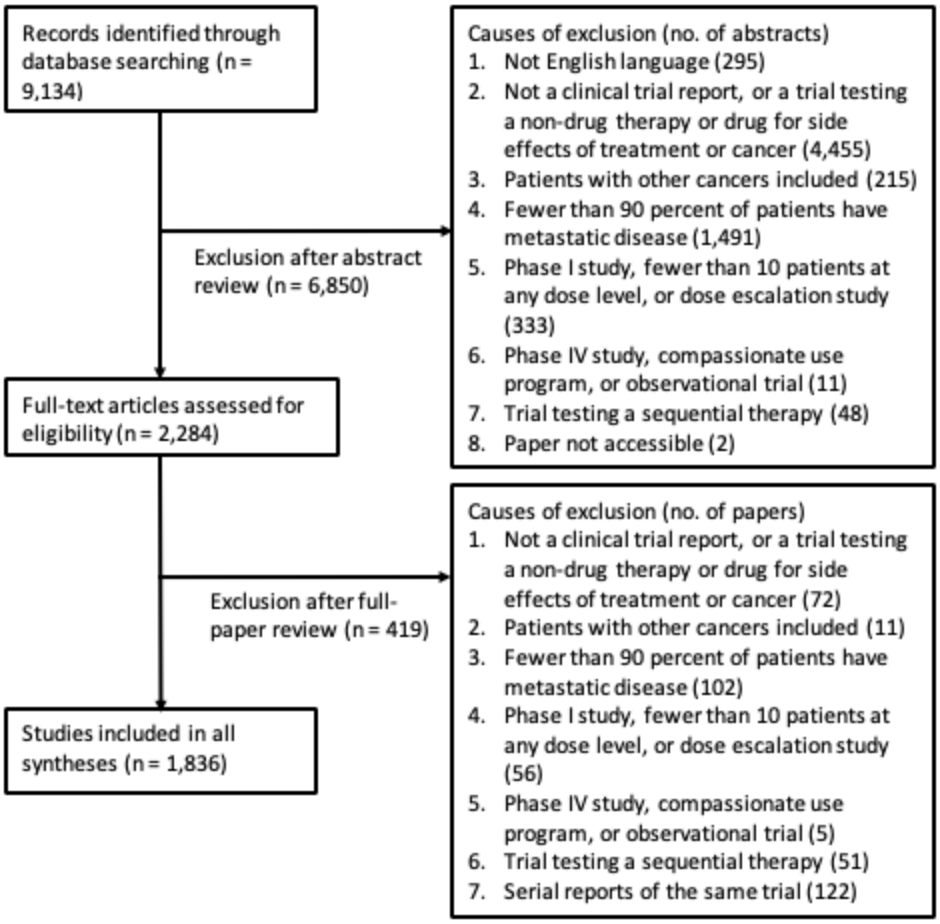
Stratum	Dose-Limiting Toxicity		Median TTP/PFS		Median Overall Survival	
	Num	I^2: Adj. Outcome	Num	I^2: Adj. Outcome	Num	I^2: Adj. Outcome
HER-2+	90	62 [40, 89]	91	42 [0, 69]	57	35 [0, 65]
HER-2+ w/o palliative anti-HER-2 therapy	51	78 [78, 90]	56	34 [0, 62]	36	16 [1, 32]
HER-2+ w/ palliative anti-HER-2 therapy	17	52 [0, 84]	15	51 [12, 80]	10	19 [0, 12]
HER-2+ w/o anti-HER-2 therapy	27	86 [84, 94]	29	19 [0, 24]	24	19 [0, 42]
HER-2+ w/ anti-HER-2 therapy	37	59 [21, 84]	33	32 [0, 56]	18	40 [23, 65]
HR+	97	48 [6, 82]	93	56 [0, 88]	62	49 [26, 80]
HR+ w/o palliative hormonal therapy	30	61 [32, 83]	36	61 [34, 89]	27	40 [0, 90]
HR+ w/o hormonal therapy	8	46 [23, 69]	13	31 [0, 54]	15	28 [0, 42]
HR+ w/ hormonal therapy	53	41 [0, 67]	52	41 [0, 81]	32	30 [0, 46]
Undirected	513	66 [41, 94]	414	38 [0, 68]	462	32 [0, 66]
Undirected w/o palliative chemo.	210	71 [60, 97]	177	38 [0, 68]	211	27 [0, 56]
Undirected w/ palliative chemo.	142	60 [12, 88]	123	34 [0, 73]	131	17 [0, 21]
Undirected w/o chemo.	45	66 [46, 92]	19	52 [22, 79]	61	21 [0, 48]
Undirected w/ chemo.	278	64 [21, 93]	231	31 [0, 62]	251	27 [0, 61]
Totals	1598	64 [32, 93]	1382	38 [0, 71]	1397	29 [0, 62]

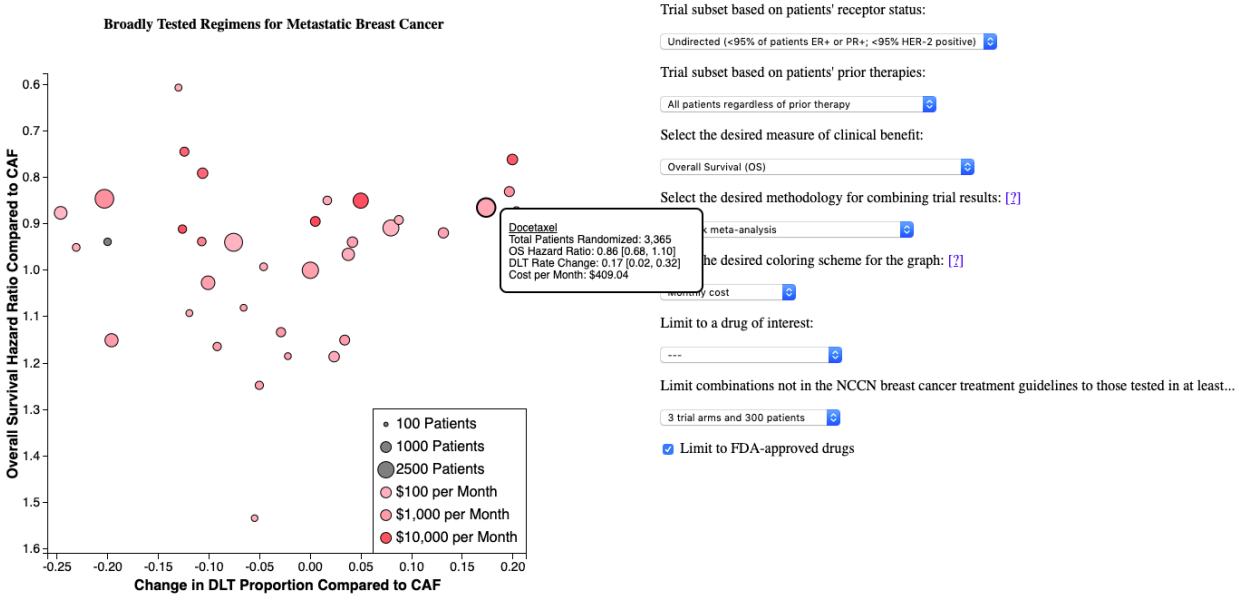
Table 3: Summary of random-effects meta-analyses performed on adjusted outcomes in each of the 14 strata. The number of meta-analyses performed are reported in the Num. Combo column, and the mean and interquartile range of I^2 values are reported.

Fig. 2 NMA estimates of OS hazard ratios and DLT risk differences of drug combinations compared to cyclophosphamide / doxorubicin / fluorouracil therapy in trials with neither HER-2-positive nor hormone receptor-positive patients, with the color of the plotted points representing the monthly drug costs of the drug combination

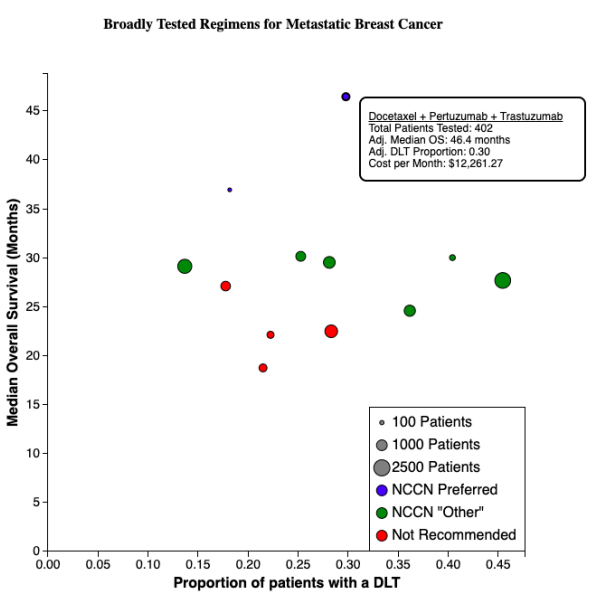
Fig. 3 Standardized estimates of the toxicity and efficacy of drug combinations in trials with HER-2-positive patients, with the color of the plotted points representing whether the drug combination is recommended in the NCCN guidelines

Fig. 4 Details of individual studies testing the docetaxel monotherapy drug combination


Records identified through database searching (n = 9,134)
Exclusion after abstract review (n = 6,850)
Causes of exclusion (no. of abstracts)
1. Not English language (295)
2. Not a clinical trial report, or a trial testing a non-drug therapy or drug for side effects of treatment or cancer (4,455)
3. Patients with other cancers included (215)
4. Fewer than 90 percent of patients have metastatic disease (1,491)
5. Phase I study, fewer than 10 patients at any dose level, or dose escalation study (333)
6. Phase IV study, compassionate use program, or observational trial (11)
7. Trial testing a sequential therapy (48)
8. Paper not accessible (2)
Full-text articles assessed for eligibility (n = 2,284)
Exclusion after full-paper review (n = 419)
Causes of exclusion (no. of papers)
1. Not a clinical trial report, or a trial testing a non-drug therapy or drug for side effects of treatment or cancer (72)
2. Patients with other cancers included (11)
3. Fewer than 90 percent of patients have metastatic disease (102)
4. Phase I study, fewer than 10 patients at any dose level, or dose escalation study (56)
5. Phase IV study, compassionate use program, or observational trial (5)
6. Trial testing a sequential therapy (51)
7. Serial reports of the same trial (122)
Studies included in all syntheses (n = 1,836)


Broadly Tested Regimens for Metastatic Breast Cancer
Overall Survival Hazard Ratio Compared to CAF
Change in DLT Proportion Compared to CAF
0.6
0.7
0.8
0.9
1.0
1.1
1.2
1.3
1.4
1.5
1.6
-0.25
-0.20
-0.15
-0.10
-0.05
0.00
0.05
0.10
0.15
0.20
Docetaxel
Total Patients Randomized: 3,365
OS Hazard Ratio: 0.86 [0.68, 1.10]
DLT Rate Change: 0.17 [0.02, 0.32]
Cost per Month: $409.04
100 Patients
1000 Patients
2500 Patients
$100 per Month
$1,000 per Month
$10,000 per Month
Trial subset based on patients' receptor status:
Undirected (<95% of patients ER+ or PR+; <95% HER-2 positive)
Trial subset based on patients' prior therapies:
All patients regardless of prior therapy
Select the desired measure of clinical benefit:
Overall Survival (OS)
Select the desired methodology for combining trial results: [?]
meta-analysis
he desired coloring scheme for the graph: [?]
Monthly cost
Limit to a drug of interest:

Limit combinations not in the NCCN breast cancer treatment guidelines to those tested in at least...
3 trial arms and 300 patients
Limit to FDA-approved drugs



Trial subset based on patients' receptor status:

HER-2 Positive (≥ 95% of patients)

Trial subset based on patients' prior therapies:

All patients regardless of prior therapy

Select the desired measure of clinical benefit:

Overall Survival (OS)

Select the desired methodology for combining trial results: [?]

Meta-analysis with demographic adjustment

Select the desired coloring scheme for the graph: [?]

NCCN guideline status

Limit to a drug of interest:

Limit combinations not in the NCCN breast cancer treatment guidelines to those tested in at least...

3 trial arms and 300 patients

☑ Limit to FDA-approved drugs

Docetaxel 60 mg/m^2 i.v. on day 1 (cycled every 21 days)

Authors	Title	Source	Year	# Patients	Patient Population	Median OS (Months)	Median PFS/TTP (Months)	DLT Proportion
S. Kim, C. Yoo, J. Ro, et al.	Combination of docetaxel and TSU-68, an oral antiangiogenic agent, in patients with metastatic breast cancer previously treated with anthracycline: Randomized phase II multicenter trial *(Trial Arm: Docetaxel)*	Investigational New Drugs	2014	39	Prior chemotherapy	26.0	8.1	0.82
N. Katsumata, T. Watanabe, H. Minami, et al.	Phase III trial of doxorubicin plus cyclophosphamide (AC), docetaxel, and alternating AC and docetaxel as front-line chemotherapy for metastatic breast cancer: Japan Clinical Oncology Group trial (JCOG9802) *(Trial Arm: D)*	Annals of Oncology	2009	147	No prior palliative chemotherapy	25.7	7.0	0.23
V. Harvey, H. Mouridsen, V. Semiglazov, et al.	Phase III Trial Comparing Three Doses of Docetaxel for Second-Line Treatment of Advanced Breast Cancer *(Trial Arm: 60 mg/m2)*	Journal of Clinical Oncology	2006	149	Prior palliative chemotherapy	10.6	3.0	0.56
T. Ishikawa, S. Shimizu, M. Inaba, et al.	A Multicenter Phase II Study of Docetaxel 60 mg/m2 as First-Line Chemotherapy in Patients with Advanced or Recurrent Breast Cancer	Breast Cancer	2004	23	No prior palliative chemotherapy	—	7.2	0.40
M. Ando, T. Watanabe, K. Nagata, et al.	Efficacy of Docetaxel 60 mg/m2 in Patients With Metastatic Breast Cancer According to the Status of Anthracycline Resistance	Journal of Clinical Oncology	2001	99	Prior palliative chemotherapy	7.4	—	0.31
I. Adachil, T. Watanabe, S. Takashima, et al.	A late phase II study of RP56976 (docetaxel) in patients with advanced or recurrent breast cancer	British Journal of Cancer	1996	72		—	3.8	—

Docetaxel 35 mg/m^2 i.v. on day 1 (cycled every 7 days)

Authors	Title	Source	Year	# Patients	Patient Population	Median OS (Months)	Median PFS/TTP (Months)	DLT Proportion
D. Pectasides, G. Papaxoinis, V. Kotoula, et al.	Expression of angiogenic markers in the peripheral blood of docetaxel-treated advanced breast cancer patients: A Hellenic Cooperative Oncology Group (HeCOG) study	Oncology Reports	2012	159	No prior palliative chemotherapy	27.7	8.8	0.21