Success Rate and Possible Causes of Failures of Phase 3 Clinical Trials in Patients with Breast Cancer: A Systematic Review

Mitsugu Ikeda, Tatsuya Ochibe and Masahiro Tohkin

Department of Regulatory Science, Graduate School of Pharmaceutical Sciences, Nagoya City University, Nagoya, Japan

Corresponding author: Masahiro Tohkin, Department of Regulatory Science, Graduate School of Pharmaceutical Sciences, Nagoya City University, Nagoya, Japan, Tel: +81 52 836 3778; E-mail: tohkin@phar.nagoya-cu.ac.jp

Received date: July 15, 2018; Accepted date: July 23, 2018; Published date: July 26, 2018

Abstract

Background: Besides the high unmet medical needs of cancer patients, the success rate of phase 3 clinical trials in this disease area is relatively low compared with other disease areas. Breast cancer is the most frequently diagnosed cancer among women. The incidence of breast cancer has increased, representing a major health problem worldwide.

Objectives: This systematic review aimed to investigate the success rate of phase 3 clinical trials in patients with breast cancer according to primary endpoints and to identify possible causes of failures to meet these endpoints.

Methods: We performed an online database search of phase 3 clinical trials in patients with breast cancer in articles published from January 2011 to June 2017, and identified “positive” (met the primary endpoint) and “negative” (failed to meet the primary endpoint) trials. The success rates were sorted by primary endpoints. Possible causes of failures to meet the primary endpoints were investigated by assessing the accuracy of pre-trial estimates.

Results: We identified 113 trials, consisting of 39 positive and 74 negative trials (overall success rate: 35%). Most of the primary endpoints (77%) were progression-related or recurrence-related. The success rates of trials assessing progression-related and recurrence-related endpoints were 39% and 17%, respectively. Progression-related and recurrence-related endpoints in the control arm showed significant improvement, compared with pre-trial estimates, which were associated with negative results.

Conclusions: The accuracy of pre-trial estimates critically influenced the success rate of phase 3 clinical trials in breast cancer patients. Although these trials need to be designed to retain the reproducibility of pre-trial estimates, the changes in diagnostic measurement and/or standard therapy from the time of study planning could provide a potential risk of underestimation of pre-trial estimates in the control arm.

Keywords: Phase 3 clinical trial; Systematic review; Breast cancer; Primary endpoint; Success rate; Causes of failures; Pre-trial estimate


Introduction

Cancer is a major public health problem worldwide and the second leading cause of death in the USA [1]. With the emergence of new anticancer agents, their efficacy and safety need to be examined in clinical trials. Randomized controlled phase 3 trials are conducted to demonstrate the superiority or non-inferiority of interventions to the standard of care, to obtain approval from regulatory agencies [2]. Besides the high unmet medical needs of cancer patients, the success rate of phase 3 clinical trials in this disease area is relatively low compared with other disease areas [3]. The success rate of phase 3 clinical trials in patients with non-small cell lung cancer (NSCLC) was 38%, based on a systematic review we conducted from articles published from January 2011 to June 2017 [4].

Breast cancer is the most frequently diagnosed cancer among women; more than 1.7 million new cases were diagnosed worldwide in 2012 [5]. Although survival rates from early breast cancer have improved substantially over the past two decades, the incidence of breast cancer has increased, representing a major health problem worldwide [6]. The biology of metastatic breast cancer has become more aggressive, developing resistance to multiple adjuvant treatment components [7]. Therefore, breast cancer remains associated with unmet medical needs and requires treatment options depending on the disease stage or treatment line. Phase 3 clinical trials select a primary endpoint to demonstrate whether a new treatment option can fill an unmet medical need, such as in patients with breast cancer [8,9].

This systematic review aimed to investigate the success rate of phase 3 clinical trials in patients with breast cancer according to primary endpoints and to identify possible causes of failures to meet these endpoints, to provide suggestions when planning phase 3 clinical trials. This systematic review was conducted in compliance with the PRISMA guidelines [10].
Methods

Search strategy

We performed an online database search through PubMed/MEDLINE, EMBASE, and the Cochrane Central Register of Controlled Trials (CENTRAL) as of December 2017 for phase 3 clinical trials in patients with breast cancer published from January 2011 to June 2017 with an available full-text paper in English, including online publications (see Appendix 1 for search terms and strategy).

One author (MI) screened titles and abstracts and selected eligible trials through a full-text review. Another author (MT) supervised and endorsed the identification of eligible trials. Any disagreements with the identification were resolved by discussion and consensus between both authors.

Criteria for collection of studies

Through the abstract review, we excluded non-phase 3 trials, trials for other diseases or tumor types, trials for non-anticancer agents or supportive care, review articles including protocol reviews, follow-up reports including sub-studies, subgroup analyses, post hoc analyses, exploratory analyses, erratum reports, pooled analysis reports, biomarker analysis reports, and any reports without primary endpoints in efficacy assessment. Eligible phase 3 clinical trials included those wherein we could identify whether the primary endpoint was met. Non-inferiority trials were included for the purpose of investigating the success rate and possible reasons for failure to meet the primary endpoint. Trials with a factorial study design, three-arm comparison trials, biosimilar trials, and formulation change trials were excluded through the full-text review.

Data extraction and analysis

All identified trials that met their primary endpoints were categorized as “positive” trials, and those that failed to meet their primary endpoints were categorized as “negative” trials. These trials were included in the analysis.

Extracted primary endpoints were categorized as follows: “response rate,” including pathological complete response (pCR) rate, clinical CR rate, overall response rate, and clinical benefit rate; “overall survival” (OS), including a co-primary endpoint with progression-free survival (PFS) or time-to-progression (TTP); “progression-related endpoint” including PFS and TTP; and “recurrence-related endpoint,” including disease-free survival (DFS), invasive DFS, recurrence-free survival (RFS), breast cancer RFS, breast cancer-free interval, event-free survival (EFS), incidence of distant metastases, and rate of invasive breast cancer events. Success rates by primary endpoints were tabulated.

We extracted actual results and pre-trial estimates of primary endpoints from full-text articles and supplemental appendixes to investigate the reasons for negative trials. We performed the paired t-test using JMP Pro 13 (SAS Institute Japan Ltd., Tokyo, Japan) to compare between actual results and pre-trial estimates.

We assessed the risk of bias in the analysis based on missing descriptions regarding the pre-estimates of primary endpoints in the trials.

Results

Study selection

A total of 640 articles were identified excluding duplications. Among these, 507 articles were excluded through abstract screening and 20 articles through full-text review. Thus, 113 phase 3 clinical trials in patients with breast cancer were included, consisting of 39 positive trials and 74 negative trials (Figure 1).

Success rates by tumor types

Success rates by tumor types are shown in Table 1. The overall success rate was 35% in this research. Of the 113 trials, 13 assessed response rate (12%), 13 assessed OS (12%), 51 assessed progression-related endpoints (45%), and 36 assessed recurrence-related endpoints (32%). Success rates for response rate (54%) and OS (46%) were higher than the overall success rate; however, the contributions to the overall success rate were limited due to the small number of trials. Since most primary endpoints were progression-related and recurrence-related, their success rates (39% and 17%, respectively) accounted for the overall success rate.
Table 1: Success rates by primary endpoints.

Response rate

Response rate was assessed in six of the 13 trials that reported negative results [11-16] (see Appendix 3 for actual results and pre-trial estimates by primary endpoints in negative trials). Overall, there was lack of reproducibility of pre-trial estimates in negative trials. Two articles described that the negative results were due to the lower-than-expected effectiveness of the experimental arm [11,12].

OS

Seven of the 13 trials reported negative results for OS [17-23] (see Appendix 3 for actual results and pre-trial estimates by primary endpoints in negative trials). Since pre-trial estimates were not found in four articles [17-20], the reproducibility of OS could not be assessed. Two articles implied that the negative results were due to the confounding by subsequent-line therapies, possibly impeding the reproducibility of pre-trial estimates of OS [18,23].

Progression-related endpoints

Available actual results and pre-trial estimates were compared in negative trials for progression-related endpoints (Figure 2). Median PFSs and TTPs were included in the analysis. From the RIBBON-1 trial [24], median PFSs for bevacizumab and placebo in both capecitabine and taxane/anthracycline cohorts were included in the analysis. In the case that the median PFS was not reached, it was not included in the analysis.

There was no significant difference in median survival time between actual results and pre-trial estimates in the experimental arm ($p=0.169, n=21$), while a significant prolongation in actual median survival times was observed, compared with pre-trial estimates, in the control arm ($p<0.001, n=22$).

Recurrence-related endpoints

Available actual results and pre-trial estimates were compared in negative trials for recurrence-related endpoints (Figure 3). Survival rates from various types of recurrence-related endpoints were included in the analysis. Survival rates in various durations (e.g., 3-year and 5-year DFS) were also included.
A significant increase in actual survival rates was observed, compared with pre-trial estimates, in both experimental (p=0.034, n=12) and control arms (p<0.001, n=14).

**Bias risk assessment**

Missing descriptions of pre-trial estimates in the articles might provide a risk of bias when assessing the relationship between accuracy of estimation and primary outcome of clinical trials. Phase 3 clinical trials without pre-trial estimates could not be included in the comparison between actual results and pre-trial estimates; however, our findings of significant improvement in progression-related and recurrence-related endpoints in the control arm are consistent with reasons for negative results with respective endpoints, described in the articles [25-40]. Therefore, the impact of missing descriptions of pre-trial estimates on the analysis is considered limited.

The definitions of individual endpoints included in the analyses for “progression-related endpoints” and “recurrence-related endpoints” were slightly different. The selection of endpoints included in the analyses for “progression-related endpoints” and “recurrence-related endpoints” might have a risk of bias. Even though the definitions of individual endpoints were slightly different, in order to investigate the accuracy of pre-trial estimates with minimum selection bias, “progression-related endpoints” included PFS or TTP, whose results were described as median survival times, and “recurrence-related endpoints” included DFS, RFS, EFS, or incidence of distant metastases, whose results were described as survival rates.

Publication bias could be negligible since the success rate of phase 3 clinical trials in this survey was consistently low, as previously reported in the oncology area [3].

**Discussion**

Clinical trials have different endpoints depending on the purpose of each trial [8]. In conventional oncology drug development, early-phase clinical trials require assessment of tumor shrinkage to identify the biological activity of a drug, and then later-phase trials commonly evaluate a clinical benefit derived from the tumor shrinkage, such as prolongation of RFS or PFS. Finally, survival benefit is confirmed.

Since breast cancer is one of the most commonly diagnosed cancers worldwide [5] and unmet medical needs exist in various stages of disease or treatment lines [6,7], substantial numbers of phase 3 clinical trials and various types of primary endpoints were expected to be collected in this research. Although sufficient numbers of phase 3 clinical trials could be identified to assess the success rate by primary endpoints and to investigate possible causes of negative results, the number of phase 3 clinical trials assessing response rate was limited. This seems reasonable because response rate is rarely used as the primary endpoint in phase 3 clinical trials. However, in 2014, the Food and Drug Administration (FDA) officially recommended pCR as the endpoint to support accelerated approval for neoadjuvant setting in high-risk early-stage breast cancer [9]. Therefore, the number of phase 3 clinical trials assessing response rate as the primary endpoint will likely increase in the next decade.

The number of phase 3 clinical trials assessing OS was also limited. Given the longer survival time in breast cancer, a longer duration is required in phase 3 clinical trials assessing OS compared with other endpoints. In this study, most primary endpoints were progression-related and recurrence-related. Therefore, the success rates of progression-related and recurrence-related endpoints accounted for the overall success rate in this research.

PFS is defined as the time from randomization until objective tumor progression or death, whereas DFS is defined as the time from
as the primary endpoints in phase 3 clinical trials and are not confounded by subsequent-line therapies. However, our research found that actual results in the control arm were significantly improved, compared with pre-trial estimates, which resulted in failures to meet the primary endpoints. The use of improved treatment options and a more favorable prognosis of patients than expected are considered possible reasons for the improvement in the control arm [25,30,32-37].

The difficulty of estimating the risk for a DFS event was discussed in the study design of the TEXT and SOFT trials in premenopausal women with endocrine-responsive early-breast cancer [40]. Assumptions to estimate the hazard relied on the results of past clinical trials and were based on treatments, standard of care, and tumor assessment tools in the past 10-20 years. When the TEXT and SOFT trials were developed, there were limited mature outcome data from trials in premenopausal women with endocrine-responsive breast cancer treated with adjuvant tamoxifen. Therefore, assumptions for estimating 5-year DFS for the TEXT and SOFT planning were based on trial data of premenopausal women who did not receive tamoxifen. A possibility of overestimation of DFS arises if updated data from current medical practice are not incorporated in the assumptions for the estimation. Additionally, if the DFS event occurs more slowly than anticipated, it could be resulting in lower statistical power and longer follow-up period.

FDA guidance mentions that “important considerations in evaluating DFS as a potential endpoint include the estimated size of the treatment effect and proven benefits of standard therapy” [8]. During study planning, it is important to adequately estimate the DFS of both test drug and control based on data on the updated standard therapy. To prevent potential bias and retain the reproducibility of pre-trial estimates, assessment of DFS, including assessment schedule in the clinical trial, should be consistent with those for assumption data. It is also important that the patient population be consistent with that in the assumption data to prevent bias due to death prior to tumor progression.

The above suggestions could be applied to any endpoints. Generally, effect size in phase 3 clinical trials is estimated based on data from phase 2 clinical trials. Considering the reproducibility of estimated effect size, when the phase 2 study design is discussed, the future phase 3 study design should be envisioned, paying attention to consistency in patient population, diagnosis method, standard therapy, and assessment measurement.

We previously conducted a systematic review of phase 3 clinical trials in patients with non-small cell lung cancer (NSCLC) using the same online database search for articles published during the same period (January 2011-June 2017) as this research [4]. Interestingly, more positive results were found in phase 3 clinical trials assessing PFS as the primary endpoint in patients with NSCLC, and no significant difference in median PFS was observed between the actual results and the pre-trial estimates in the control arm. The reproducibility of the pre-trial estimates of controlled PFS might vary by tumor type. Thus, the accuracy of the pre-trial estimates of controlled PFS of phase 3 clinical trials in breast cancer patients might need more careful consideration than that for patients with NSCLC.

Our study has some limitations. The sample sizes of trials assessing response rate and OS were limited. Lack of capability of the tested drugs to improve the response rate and confounding by subsequent-line therapies for improving the OS were considered as possible causes of negative results with respective primary endpoints; however, the reasons for these findings could not be investigated in this study. One possible reason might be false-positive results in phase 2 clinical trials. Hence, further investigation of previous phase 2 clinical trials that served as basis for planning phase 3 clinical trials might provide clarification.

Progression-related and recurrence-related endpoints might differ depending on disease status, including hormone or human epidermal growth factor receptor 2 status, and/or experimental drugs with different modes of action. Due to the limited sample size, the impact of these factors could not be considered for our findings. However, the trend found in this study has implications in terms of the importance of accuracy of pre-trial estimates. Moreover, due to the limited number of positive trials, we could not assess the difference in accuracy of pre-trial estimates between positive and negative trials.

Another limitation is the duration of data collection in the systematic review, which might provide different success rates. More evident tumor biology could target a more enriched patient population in clinical trials. Novel targeted therapies might bring higher success rates of phase 3 clinical trials in the future. Furthermore, accumulated data from a well-organized clinical study design may contribute to more-accurate pre-trial estimates of the primary endpoints in the future, which may also result in higher success rates of phase 3 clinical trials.

Conclusion

The overall success rate of phase 3 clinical trials in breast cancer patients was 35% in this study. Most primary endpoints included in the analysis were progression-related or recurrence-related, which accounted for the overall success rate. There was significant improvement in progression-related or recurrence-related endpoints in the control arm, which resulted in negative outcomes of phase 3 clinical trials. The changes in diagnostic measurement and/or standard therapy from the time of study planning could lead to potential risk of underestimation of progression-related or recurrence-related endpoints in the control arm. Therefore, the patient population, diagnosis method, standard therapy, and assessment measurement require consistency during study planning to retain the reproducibility of pre-trial estimates.

Conflict of Interest

Mitsugu Ikeda is an employee of Daiichi Sankyo Co., Ltd. The authors declare no other conflicts of interest. This study was not funded by any foundation or company.

References


