Progression-free Survival With First-line Endocrine-based Therapies Among Postmenopausal Women With HR+/HER2– Metastatic Breast Cancer: A Network Meta-analysis

Rajeev Ayyagari, PhD1; Derek Tang, PhD2; Oscar Patterson-Lomba, PhD1; Zhou Zhou, MSc1; Jipan Xie, PhD1; David Chandiwana, MSc2; Anand A. Dalal, PhD, MBA2; and Polly Ann Niravath, MD3

1Analysis Group, Inc, Boston, Massachusetts; 2Novartis Pharmaceuticals Corporation, East Hanover, New Jersey; and 3Methodist Hospital, Houston, Texas

ABSTRACT

Purpose: The comparative efficacy of endocrine-based therapies (ETs) for hormone receptor–positive/human epidermal growth factor receptor 2–negative (HR+/HER2–) metastatic breast cancer (mBC) is not well characterized. This network meta-analysis (NMA) synthesized available evidence on progression-free survival (PFS) with first-line ETs for postmenopausal HR+/HER2– mBC.

Methods: A systematic literature review identified randomized controlled trials of first-line ETs. Pairwise hazard ratios and 95% credible intervals (CrIs) were obtained via a Bayesian NMA model. Subgroup NMAs were conducted among late progressors (disease-free interval ≥12 months from completion of [neo]adjuvant therapy with letrozole or anastrozole at the time of randomization) and de novo patients, defined as patients whose initial BC diagnosis is mBC.

Findings: Five trials and 5 regimens (ribociclib + an aromatase inhibitor [AI] [LEE + AI], palbociclib + AI [Pal + AI], fulvestrant 250 mg + AI [Ful250 + AI], fulvestrant 500 mg [Ful500], and AI) were selected. LEE + AI, Pal + AI, Ful250 + AI, and Ful500 had significantly longer PFS versus AI (95% CrI upper-bound ≤1). LEE + AI had a 30% and 29%, and Pal + AI had a 31% and 30%, reduced hazard of progression or death versus Ful500, respectively, but were not statistically significant. In both subgroup analyses, all therapies had significantly longer PFS compared with AI.

Implications: Pal + AI, LEE + AI, Ful250 + AI, or Ful500 as first-line treatment for HR+/HER2– mBC had longer PFS than AI alone. Given the lack of head-to-head clinical trials comparing the efficacy of recently approved first-line ETs for HR+/HER2– mBC, these results have important clinical implications for the treatment of HR+/HER2– mBC in the first-line setting. (Clin Ther. 2018;40:628–639) © 2018 Elsevier HS Journals, Inc. All rights reserved.

Key words: endocrine therapy, HR positive/HER2 negative, metastatic breast cancer, network meta-analysis, progression-free survival, targeted therapy.

INTRODUCTION

Breast cancer (BC) is the most common type of cancer in women worldwide.1 In the United States (US), >252,000 new cases and 40,000 BC-related deaths are expected in 2017.2 According to a US study,3

Accepted for publication March 7, 2018. https://doi.org/10.1016/j.clinthera.2018.03.004 0149-2918/$ - see front matter © 2018 Elsevier HS Journals, Inc. All rights reserved.
~70% of all BCs test positive for estrogen or progesterone hormone receptor (HR) and negative for human epidermal growth factor receptor-2 (HER2), referred to as HR+/HER2− BC. The likelihood of being diagnosed with HR+/HER2− BC seems to be positively associated with age, with postmenopausal women at particularly high risk due to hormonal factors.4,5

Although 90% to 95% of all BC cases are detected at an early stage, ~30% eventually become metastatic (mBC),6 resulting in a 5-year survival rate of 26%.2 Patients with mBC have been shown to have a significantly reduced quality of life compared with the general population, particularly in terms of physical, emotional, and social functioning.7 As the disease progresses and symptoms worsen, loss of function becomes increasingly debilitating, both physically and emotionally. This scenario translates into a substantial economic burden for both patients and society, as evidenced by the high health care costs associated with BC, which are estimated to represent as much as 20% of the total costs of cancer care in the US.8

Targeted therapies (TTs) are poised to revolutionize the treatment of mBC as, in combination with endocrine therapy, they have been shown to inhibit cancer progression while avoiding endocrine resistance and offering manageable safety profiles.9,10 Indeed, in the case of postmenopausal women with not immediately life-threatening HR+/HER2− mBC, the National Comprehensive Cancer Network treatment guidelines11 currently recommend the use of endocrine-based therapy (ET). These recommendations include third-generation nonsteroidal aromatase inhibitors (AIs) such as anastrozole and letrozole, the steroidal AI exemestane, and the selective estrogen receptor downregulator fulvestrant,12 with letrozole,13 and endocrine therapy in combination with TT including everolimus+exemestane,14 palbociclib + letrozole,15,16 and ribociclib + letrozole.19 If ET fails, chemotherapy is recommended when symptomatic visceral disease is present or the cancer is either rapidly progressing or immediately life-threatening.11 However, chemotherapy is often accompanied by serious adverse events of grade 3/4, further impairing a patient’s health-related quality of life.20

Head-to-head comparison data on the relative efficacy of first-line therapies for HR+/HER2− BC are lacking. Previous studies have compared these therapies by using network meta-analyses (NMAs).21–23 However these NMAs were not limited to HR+/HER2− patients and did not account for single lines of therapy, a critical predictor of outcome, but, instead, jointly analyzed first- and second-line settings. Moreover, none of them compared the efficacy of ribociclib versus that of Endocrine-based therapies (either as monotherapy or in combination with TTs or other ETs) in patients with HR+/HER2− mBC in the first-line setting given that the results of the MONALEESA-2 (Mammary Oncology Assessment of LEE011’s [Ribociclib’s] Efficacy and Safety) trial,19 which compared ribociclib + letrozole versus placebo + letrozole in previously untreated HR+/HER2− BC, were only recently made available. To fill this knowledge gap, using an NMA, the goal of the present study was to quantitatively synthesize the available evidence on progression-free survival (PFS) associated with endocrine therapies as first-line treatments for HR+/HER2− mBC among postmenopausal women, and evaluate the efficacy of these agents in pre-identified patient populations in an effort to identify subsets of patients most likely to benefit from novel TTs as well as to explore comparative efficacy in more homogeneous settings.

MATERIALS AND METHODS
Systematic Literature Review
A systematic literature review was conducted on June 7, 2016, to identify published randomized controlled trials with efficacy data of endocrine therapy (ie, letrozole, anastrozole, exemestane, tamoxifen, fulvestrant) or TT (ie, palbociclib, everolimus, ribociclib, abemaciclib), either as monotherapy or as part of a combination therapy, for the treatment of women with HR+/HER2− mBC. The search was performed in the following databases: Medline, EMBASE, Cochrane databases (Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, and Database of Abstracts of Reviews of Effects), and conference proceedings from 2013 to 2016 (American Society for Clinical Oncology, American Association for Cancer Research, American Society for Clinical Oncology Breast Cancer Symposium, San Antonio Breast Cancer Symposium, European CanCer Organisation, European Breast Cancer Conference, and European Society of Medical Oncology). Search terms were constructed for the following 3 domains: (1) disease (ie, mBC); (2) treatments (ie,
ET); and (3) study type (ie, *British Medical Journal* filter for randomized controlled trials). To decrease heterogeneity in the patient populations, the search was restricted to 2007 and onward given that the standardization of HER2 testing was developed in 2007 by the American Society of Clinical Oncology/College of American Pathologists.24 The systematic literature review was designed, performed, and reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis guidelines.25 Study quality was assessed by using the Cochrane Collaboration’s tool for assessing risk of bias.26

The inclusion and exclusion criteria were defined a priori and are listed in Table I. Based on these criteria, the literature screening was conducted in 2 steps. As a first step (level I screening), all titles/abstracts for the identified studies were reviewed based on the inclusion and exclusion criteria. For the abstracts that passed level I screening, the corresponding full-text articles

### Table I. Inclusion and exclusion criteria.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Inclusion Criteria</th>
<th>Exclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>Postmenopausal women with HR+/HER2− mBC in first-line therapy setting</td>
<td>Non-HR+/HER2− subtype or study does not report PFS outcomes separately for this subtype</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Not mBC or study includes a mixed population but does not report PFS results separately for mBC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Trials primarily focused on second-line or later lines of therapy</td>
</tr>
<tr>
<td>Interventions</td>
<td>At least 1 of the following therapies, either as monotherapy or as part of a combination therapy</td>
<td>No drug of interest included</td>
</tr>
<tr>
<td>and comparators</td>
<td>Endocrine monotherapy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Letrozole</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Anastrozole</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Exemestane</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Tamoxifen</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Fulvestrant</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Targeted therapy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Palbociclib</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Everolimus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Ribociclib</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Abemaciclib</td>
<td></td>
</tr>
<tr>
<td>Outcomes</td>
<td>PFS</td>
<td>No PFS outcome reported</td>
</tr>
<tr>
<td>Study design</td>
<td>RCT</td>
<td>Observational studies</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Single-arm trials</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Case reports</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Editorials and opinion pieces</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reviews</td>
</tr>
</tbody>
</table>

HR+/HER2− = hormone receptor-positive/human epidermal growth factor receptor 2-negative; mBC = metastatic breast cancer; PFS = progression-free survival; RCT = randomized controlled trial.
were retrieved for further review (level II screening). In addition, any relevant systematic reviews or meta-analyses published during the past 3 years were reviewed to identify additional studies among the articles they cited.

The reasons for exclusion were recorded for all the excluded articles. The number of studies excluded for different reasons were summarized at each level of screening.

Study-level information (eg, trial acronym, population), PFS outcomes (eg, hazard ratios and their corresponding SEs), and the reported baseline characteristics were extracted from the selected studies. To ensure accuracy of the study selection and data extraction, literature screening and data extraction were performed by 2 researchers independently. A third researcher was consulted to reach a consensus in cases of disagreement on inclusion/exclusion decisions or extracted data.

### Network Meta-analysis

The relative treatment effects in patients receiving first-line treatment for HR+/HER2– mBC were estimated via a series of NMAs. To determine trial inclusion in the networks of evidence, trials that passed level II screening were further screened for the availability of PFS data and the explicit reporting of hazard ratios and SEs or 95% CIs (from which SEs were estimated).

To enable the formation of evidence networks, and based on clinical input, the AIs anastrozole, letrozole, and exemestane were pooled together into a single arm in the core analysis. The resulting efficacy of the pooled AI arm is effectively a weighted average of the efficacies of all the arms pooled, weighted according to their respective SEs reported in each separate AI arm.

In addition to the core analysis on patients receiving first-line treatment for HR+/HER2– mBC, 2 subgroups were studied separately. First, we studied the subgroup of late progressors, defined as patients with a disease-free interval ≥12 months from completion of (neo) adjuvant therapy with letrozole or anastrozole at the time of randomization. Second, we investigated the subgroup of "de novo" patients. Nominally, de novo patients were considered those whose initial BC diagnosis is mBC. For each of the 2 subgroups, a network diagram was created, showing the sources of evidence available for the analysis.

For each of the 3 populations of interest, an NMA model was fitted by using the log of the PFS hazard ratio for each treatment comparison in each study via fixed effects and random effects models to account for different assumptions regarding heterogeneity of treatment effects. The geometry of the networks was such that, in most cases, there was 1 study per treatment comparison, leading to unreliable parameter estimates and wide credible intervals (CrIs) for the random effects models. Therefore, the results presented here are from the fixed effects model.

Based on the most recent guidance from the National Institute for Health and Care Excellence Decision Support Unit and the International Society for Pharmacoeconomics and Outcomes Research good practices for indirect comparison, we employed a Bayesian approach using a normal likelihood model with linear link. The Bayesian approach provides posterior distributions for parameters of interest that allow for both statistical estimation and inference (summary measures such as mean and Crls) and a framework for probabilistic decision-making under uncertainty (eg, estimates of the probability that each treatment is superior). Noninformative previous distributions were used as parameters in the NMA to avoid artificially biasing results and ensuring maximal objectivity of the results.

An estimate of treatment effect comparing each pair of treatments was generated using the NMA in the odds ratio scale using the pooled reference-arm risk in the Bayesian model and converting the odds ratio into a relative risk, in accordance with National Institute for Health and Care Excellence guidelines. Ten thousand iterations were used as burn-in, following with an additional 20,000 iterations run within the Bayesian NMA to generate the posterior distributions for the treatment effects for each outcome of interest. The uncertainty of the estimates was summarized by using 95% Bayesian Crls. In addition, the probability that each drug is ranked as having the highest efficacy, or the second highest efficacy, and so forth, among all treatments was derived from the posterior distributions of each treatment efficacy.

Heterogeneity in the evidence network was assessed by comparing the deviance information criterion for the fixed and random effects models. A better fit...
(lower deviance information criterion) for the fixed effects model was taken to indicate that between-study heterogeneity in the effect size is low across the entire network. In addition, the $I^2$ and Cochran’s Q statistics were used to assess the heterogeneity of outcomes in PALOMA (Palbociclib: Ongoing Trials in the Management of Breast Cancer)-15 and PALOMA-2,16 as these 2 trials connected Pal + AI and AI, and represented the only link in the network with >1 trial. A significant Q statistic, or one that is large in relation to degrees of freedom, can imply heterogeneity. $I^2$ measures the degree of heterogeneity and represents the proportion of total variability among the effect sizes from different trials that is not explained by sampling error. Transitivity was assessed by comparing baseline characteristics across arms and trials. All of the analyses were implemented by using statistical software R (R Foundation for Statistical Computing, Vienna, Austria)17 and JAGS (a software package for implementing Markov Chain Monte Carlo techniques).18

RESULTS

Study Selection

For the core analysis in the first-line setting, 5 randomized controlled trials, namely PALOMA-1,15 PALOMA-2,16 MONALEESA-2,19 FALCON (Fulvestrant and Anastrozole Compared in Hormonal Therapy-Naive Advanced Breast Cancer),12 and Mehta et al,13 met the screening criteria and were included in the analyses (Figure 1). These 5 randomized controlled trials included 5 regimens: palbociclib + AI (Pal + AI), ribociclib + AI (LEE + AI), fulvestrant 250 mg + AI (Ful250 + AI), fulvestrant 500 mg (Ful500), and AI. The sample sizes in each arm ranged from 81 to 444 patients.

Subgroup analysis among late progressors included 2 clinical trials (PALOMA-216 and MONALEESA-219) and 3 regimens (Pal + AI, LEE + AI, and AI). In the subgroup analysis considering only trials reporting results for "de novo" patients (patients whose initial BC diagnosis is mBC), 4 trials (PALOMA-1,15 PALOMA-2,16 MONALEESA-2,19 and FALCON12) and 4 regimens (Pal + AI, LEE + AI, Ful500, and AI) were eligible. For the PALOMA-115 and PALOMA-216 trials, patients with de novo metastases were considered de novo patients. For MONALEESA-2,19 patients with newly diagnosed mBC were considered de novo patients. In the FALCON trial,12 patients with no previous exposure to chemotherapies were considered de novo patients.

The evidence networks from the 3 populations investigated (core, late progressors, and de novo patients) are presented in Figure 2.
Trial Characteristics

Baseline demographic characteristics and disease status of the patients who participated in the selected trials are presented in Table II. Baseline prior treatments are presented in Supplemental Table I (see the online version at https://doi.org/10.1016/j.clinthera.2018.03.004).

Although there was between-trial variability in the baseline characteristics of the selected trials, the distributions of these characteristics are largely similar clinically. Thus, the transitivity (similarity) assumption is not violated in the analyses presented in this study. Table III summarizes the trial-level treatment effects of the selected trials.

Treatment Comparisons for Core Analysis

LEE + AI, Pal + AI, Ful250 + AI, and Ful500 had significantly longer PFS versus AI (95% CrI upper bound ≤1). The median hazard ratio compared with AI was 0.57 (95% CrI, 0.46, 0.71) and 0.56 (95% CrI, 0.46, 0.68) for LEE + AI and Pal + AI, respectively. The median hazard ratio comparing LEE + AI to Pal + AI was 1.02 (95% CrI, 0.76, 1.36) (Table IV). LEE + AI had a 46% probability of being the most efficacious treatment, whereas Pal + AI had a 54% probability (Figure 3).

Treatment Comparisons for Late Progressors Analysis

Among late progressors (patients with disease-free interval ≥12 months from completion of [neo] adjuvant therapy), LEE + AI and Pal + AI also had significantly longer PFS than AI (95% CrI upper bound ≤1). The median hazard ratio compared with AI was 0.56 (95% CrI, 0.43, 0.72) and 0.58 (95% CrI, 0.46, 0.73) for LEE + AI and Pal + AI, respectively. The median hazard ratio comparing LEE + AI versus Pal + AI was 0.96 (95% CrI, 0.68, 1.34) (see Supplemental Table II in the online version at https://doi.org/10.1016/j.clinthera.2018.03.004). LEE + AI had a 60% probability of being the most efficacious treatment versus a 40% probability for Pal + AI (see Supplemental Figure 1 in the online version at https://doi.org/10.1016/j.clinthera.2018.03.004).

Treatment Comparisons for De Novo Analysis

In the subgroup analysis considering only trials reporting results for de novo patients, LEE + AI and Pal + AI again exhibited significantly better efficacy than other treatments in terms of PFS. The median hazard ratio compared with AI was 0.45 (0.27, 0.76) for LEE + AI, 0.54 (95% CrI, 0.39, 0.74) for Pal + AI, and 0.75 (95% CrI, 0.58, 0.97) for Ful500. The median hazard ratio comparing LEE + AI versus Pal + AI was 0.84 (95% CrI, 0.46, 1.55) (see Supplemental Table III in the online version at https://doi.org/10.1016/j.clinthera.2018.03.004). LEE + AI and Pal + AI had a 71% and 29% probability of being the most efficacious treatments, respectively (see Supplemental Figure 2 in the online version at https://doi.org/10.1016/j.clinthera.2018.03.004).

Heterogeneity Assessment

Based on National Institute for Health and Care Excellence guidelines, outcome heterogeneity was assessed by using \( I^2 \) and Cochran’s Q measures for any pair of drugs that were compared in ≥2 trials. No evidence of heterogeneity was found in the networks except in the de novo subgroup analysis (\( I^2 = 0.78; Q P \) value = 0.035). This finding is expected because in...
Table II. Baseline demographic characteristics and disease status. The values are given as no. (%) unless otherwise indicated.

<table>
<thead>
<tr>
<th>Trial</th>
<th>Anastrozole + Fulvestrant</th>
<th>Anastrozole</th>
<th>Letrozole + Palbociclib</th>
<th>Letrozole</th>
<th>Letrozole + Palbociclib</th>
<th>Letrozole + LEE</th>
<th>Letrozole</th>
<th>Fulvestrant</th>
<th>Anastrozole</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, median (range), y</td>
<td>NR</td>
<td>NR</td>
<td>63 (54–71)</td>
<td>64 (56–70)</td>
<td>62 (30–89)</td>
<td>61 (28–88)</td>
<td>63 (29–88)</td>
<td>62 (23–91)</td>
<td>64 (38–87)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>344 (77.5)</td>
<td>172 (77.5)</td>
<td>269 (80.5)</td>
<td>280 (83.8)</td>
<td>175 (76.1)</td>
<td>174 (75.0)</td>
</tr>
<tr>
<td>Asian</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>65 (14.6)</td>
<td>30 (13.5)</td>
<td>28 (8.4)</td>
<td>23 (6.9)</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Black</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>8 (1.8)</td>
<td>3 (1.4)</td>
<td>10 (3.0)</td>
<td>7 (2.1)</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Native American</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>0</td>
<td>0</td>
<td>1 (0.3)</td>
<td>0</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Pacific Islander</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>0</td>
<td>0</td>
<td>1 (0.3)</td>
<td>0</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Other</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>27 (6.1)</td>
<td>17 (7.7)</td>
<td>12 (3.6)</td>
<td>8 (2.4)</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Unavailable</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>0</td>
<td>0</td>
<td>13 (3.9)</td>
<td>16 (4.8)</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Hormone receptor status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ER+, PR+</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>269 (80.5)</td>
<td>277 (82.9)</td>
<td>175 (76.1)</td>
<td>179 (77.2)</td>
</tr>
<tr>
<td>ER+, PR-</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>44 (19.1)</td>
<td>43 (18.5)</td>
</tr>
<tr>
<td>ER+, PR unknown</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>10 (4.3)</td>
<td>7 (3.0)</td>
</tr>
<tr>
<td>ER- or unknown, PR+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ER unknown, PR unknown</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECOG 0</td>
<td>NR</td>
<td>NR</td>
<td>46 (55.0)</td>
<td>45 (56.0)</td>
<td>257 (57.9)</td>
<td>102 (45.9)</td>
<td>205 (61.4)</td>
<td>202 (60.5)</td>
<td>NR</td>
</tr>
<tr>
<td>ECOG 1</td>
<td>NR</td>
<td>NR</td>
<td>38 (45.0)</td>
<td>36 (44.0)</td>
<td>178 (40.1)</td>
<td>117 (52.7)</td>
<td>129 (38.6)</td>
<td>132 (39.5)</td>
<td>NR</td>
</tr>
<tr>
<td>ECOG 2</td>
<td>NR</td>
<td>NR</td>
<td>0</td>
<td>0</td>
<td>9 (2.0)</td>
<td>3 (1.4)</td>
<td>0</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>ECOG &gt; 2</td>
<td>NR</td>
<td>NR</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Unavailable</td>
<td>NR</td>
<td>NR</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Disease stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Locally advanced†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metastatic‡</td>
<td>NR</td>
<td>NR</td>
<td>82 (98.0)</td>
<td>80 (99.0)</td>
<td>138 (31.1)</td>
<td>72 (32.4)</td>
<td>333 (99.7)</td>
<td>331 (99.1)</td>
<td>202 (87.8)</td>
</tr>
<tr>
<td>Metastatic site of cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bone</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>246 (73.7)</td>
<td>244 (73.1)</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Bone only</td>
<td>NR</td>
<td>NR</td>
<td>17 (20.0)</td>
<td>12 (15.0)</td>
<td>103 (23.2)</td>
<td>48 (21.6)</td>
<td>69 (20.7)</td>
<td>78 (23.4)</td>
<td>NR</td>
</tr>
<tr>
<td>Visceral</td>
<td>NR</td>
<td>NR</td>
<td>37 (44.0)</td>
<td>43 (53.0)</td>
<td>214 (48.2)</td>
<td>110 (49.5)</td>
<td>197 (59.0)</td>
<td>196 (58.7)</td>
<td>135 (58.7)</td>
</tr>
<tr>
<td>Measurable disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NR</td>
<td>65 (77.4)</td>
<td>66 (81.5)</td>
<td>338 (76.1)</td>
<td>171 (77.0)</td>
<td>256 (76.6)</td>
<td>245 (73.4)</td>
<td>193 (83.9)</td>
<td>196 (84.5)</td>
<td></td>
</tr>
<tr>
<td>Disease-free interval</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>De novo</td>
<td>NR</td>
<td>NR</td>
<td>44 (52.0)</td>
<td>37 (46.0)</td>
<td>167 (37.6)</td>
<td>81 (36.5)</td>
<td>114 (34.1)</td>
<td>113 (33.8)</td>
<td>NR</td>
</tr>
<tr>
<td>≤ 2 mo</td>
<td>NR</td>
<td>NR</td>
<td>59 (70.0)</td>
<td>51 (63.0)</td>
<td>99 (22.3)</td>
<td>48 (21.6)</td>
<td>4 (1.2)</td>
<td>10 (3.0)</td>
<td>NR</td>
</tr>
<tr>
<td>&gt; 12 mo</td>
<td>NR</td>
<td>NR</td>
<td>25 (30.0)</td>
<td>30 (37.0)</td>
<td>178 (40.1)</td>
<td>93 (41.9)</td>
<td>216 (64.7)</td>
<td>210 (62.9)</td>
<td>NR</td>
</tr>
<tr>
<td>Unknown</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>0</td>
<td>1 (0.3)</td>
<td>NR</td>
<td>NR</td>
</tr>
</tbody>
</table>

ECOG = Eastern Cooperative Oncology Group; ER = estrogen receptor; LEE = ribociclib; PR = progesterone receptor; NR = not reported.
*The overall population in the 2012 study by Mehta et al is hormone receptor-positive, which is broader than hormone receptor-positive/human epidermal growth factor receptor 2-negative. The human epidermal growth factor receptor 2-negative subgroup results were used.
†Stage III cancer was considered to be locally advanced disease.
‡Stage IV cancer was considered metastatic.
most of the evidence networks presented, there was only 1 trial per link. The only 2 treatments connected by 2 trials (PALOMA-1 \(^{15}\) and PALOMA-2 \(^{16}\)) were Pal + AI and AI. In all of the analyses, the deviance information criterion value for the fixed effects model was lower than that of the random effects model, except in the case of the de novo subgroup analysis. In all cases, the fixed effects models provided more stable results and converged posterior densities. Therefore, we presented the results from the fixed effect model.

**Quality Assessment Results of Studies Included in the Indirect Treatment Comparison**

Overall, the risk of bias for the included trials was low to moderate, with some trials not reporting some information such as concealment of allocation and blinding of care providers and participants. No adjustments were made to the analysis (see Supplemental Table IV in the online version at https://doi.org/10.1016/j.clinthera.2018.03.004).

**DISCUSSION**

With the rapid changes in the treatment landscape for postmenopausal women with HR+/-HER2–mBC, there is a need for patients, clinicians, and payers to assess the relative benefits and risks of the emerging and currently available endocrine-based therapies. Such an assessment should take a comprehensive view of the available evidence and include considerations of the relative efficacy of the treatments. However, this type of assessment is difficult to make based on the available clinical trials of first-line treatments in this population due to limitations of the available evidence. Although each of the newer endocrine-based therapies has been included in clinical trials, each trial only includes head-to-head comparisons between 2 treatments, and no trial compares newer therapies such as Pal + AI and LEE + AI. Hence, there is a need to use indirect comparison methods that can synthesize information on multiple treatment alternatives in different randomized trials to produce comparisons of the efficacy of these treatments. The use of an NMA approach allows synthesis of a large body of evidence while retaining the benefits of randomization within each clinical trial. These methods are being rapidly adopted across a wide range of health research areas to synthesize direct and indirect evidence from different treatment comparisons and to inform economic decision models. This evidence is particularly
attractive for health technology appraisals and for clinicians, whose main interest is to identify the most efficacious and cost-effective treatment.31,32

The findings of the NMA-based evidence synthesis are consistent with those in the individual randomized trials. However, the analyses herein performed also provide, to our knowledge, the first estimation of the comparative efficacy between Pal + AI and LEE + AI against FUL and FUL + AI in the first-line setting, indicating a clear clinical benefit of the former therapies over the latter ones in terms of PFS. In the core analysis, we found that Pal + AI, LEE + AI, FUL + AI, and FUL had significantly longer PFS versus AI, with the greatest benefit observed for Pal + AI and LEE + AI. Moreover, LEE + AI and Pal + AI both had a 30% and 29% and Pal + AI had a 31% and 30% reduced hazard of progression or death versus Ful250 + AI and Ful500 (95% CrI upper bound \(\leq 1\)), respectively. A recent guideline recommendation from the American Society of Clinical Oncology for the treatment of mBC suggests the use of combination hormone therapy plus AI and Ful500 + AI in the first-line setting.33 These recommendations can be further improved in light of our findings by pointing out that targeted therapies combined with AI can substantially improve PFS compared with Ful500 and Ful250 + AI.

Among late progressors, LEE + AI had a 4% reduced hazard of progression or death versus Pal + AI, although this reduction was not statistically significant. LEE + AI had a higher probability of being the most efficacious (60%) followed by Pal + AI (29%).

The analysis in de novo patients revealed that LEE + AI and Pal + AI had substantial reductions in the hazard of progression or death compared with Ful500 (40% and 29%, respectively). LEE + AI had a higher probability of being the most efficacious among the compared treatments (71%) followed by Pal + AI (29%).

Altogether, these results suggest a directional benefit in terms of PFS for the 2 cyclin-dependent kinase inhibitor therapies, LEE + AI and Pal + AI; however, no statistically significant differences were found. Future research should consider targeting clinically meaningful subgroups to understand the comparative efficacy of cyclin-dependent kinase inhibitor therapies.

Although previous reviews and meta-analyses of treatments in mBC have been conducted, many of these studies focused on populations different from

---

**Table IV. Pairwise treatment comparison, core analysis. Results are given as median and 95% credible intervals of hazard ratio (column versus row).**

<table>
<thead>
<tr>
<th>Variable</th>
<th>AI</th>
<th>Ful250 + AI</th>
<th>Ful500</th>
<th>LEE + AI</th>
<th>Pal + AI</th>
</tr>
</thead>
<tbody>
<tr>
<td>AI</td>
<td>1</td>
<td>0.81 (0.67, 0.98)</td>
<td>0.80 (0.63, 1.00)</td>
<td>0.57 (0.46, 0.71)</td>
<td>0.56 (0.46, 0.68)</td>
</tr>
<tr>
<td>Ful250 + AI</td>
<td>1.23 (1.02, 1.49)</td>
<td>1</td>
<td>0.98 (0.73, 1.32)</td>
<td>0.70 (0.53, 0.94)</td>
<td>0.69 (0.53, 0.91)</td>
</tr>
<tr>
<td>Ful500</td>
<td>1.25 (1.00, 1.58)</td>
<td>1.02 (0.76, 1.37)</td>
<td>1</td>
<td>0.71 (0.52, 0.98)</td>
<td>0.70 (0.52, 0.95)</td>
</tr>
<tr>
<td>LEE + AI</td>
<td>1.76 (1.42, 2.18)</td>
<td>1.43 (1.07, 1.90)</td>
<td>1.40 (1.02, 1.91)</td>
<td>1</td>
<td>0.98 (0.74, 1.32)</td>
</tr>
<tr>
<td>Pal + AI</td>
<td>1.79 (1.46, 2.18)</td>
<td>1.45 (1.10, 1.90)</td>
<td>1.43 (1.05, 1.92)</td>
<td>1.02 (0.76, 1.36)</td>
<td>1</td>
</tr>
</tbody>
</table>

AI = aromatase inhibitor; Ful250 = fulvestrant 250 mg; Ful500 = fulvestrant 500 mg; LEE = ribociclib; Pal = palbociclib.

---

**Figure 3.** Ranking probabilities according to treatment, core analysis. AI = aromatase inhibitor; Ful250 = fulvestrant 250 mg; Ful500 = fulvestrant 500 mg; LEE = ribociclib; Pal = palbociclib.
the current study. In addition, the present analysis includes recently published data that were not included in previous reviews. Some closely related earlier review articles include those by Generali et al,21 Chirila et al,22 and Wilson et al.23 Generali et al21 conducted an NMA of randomized controlled trials focusing on the efficacy (PFS and time to progression) and toxicity of everolimus plus exemestane versus chemotherapies among HR+ mBC patients in both first- and second-line settings. The study found that the combination of everolimus plus exemestane as first- or second-line therapy for HR+ mBC was more efficacious than several chemotherapy regimens reported in the literature. The population of the trials in this study were not limited to HR+/HER2− patients only, and it analyzed the first- and second-line settings jointly. Chirila et al22 performed an NMA was associated with significant improvement in PFS and ranked more favorably than all 3 comparators (letrozole, tamoxifen, and anastrozole). Wilson et al23 conducted an NMA comparing palbociclib with chemotherapy agents for the treatment of postmenopausal women with HR+/HER2− mBC and found that in both the first- and second-line settings, palbociclib plus letrozole was associated with significantly longer PFS than all 3 comparators (letrozole, tamoxifen, and anastrozole). Wilson et al23 conducted an NMA comparing palbociclib with chemotherapy agents for the treatment of postmenopausal women with HR+/HER2− mBC and found that in both the first- and second-line settings, palbociclib plus letrozole showed significant improvements in PFS and ranked more favorably than all chemotherapy comparators. As in the study by Generali et al,21 the study by Wilson et al also included trials with populations who were not exclusively HR+ HER2− and used second-line trials to ensure the connectivity of the first-line network.

Due to the limited data available for this population, the present study had some limitations. First, the total number of studies identified was low (5 randomized controlled trials), which had several effects on the analyses. The low number of studies meant that the findings are based on the experience of a relatively limited number of patients, although this factor is offset, to some extent, by the fact that some of the studies included a large number of patients. In an NMA, factors such as differences in patient characteristics and study design among the included trials may bias the assessment of the comparative efficacy of the treatments considered in the trials. The effect that these factors have on the results of an NMA can generally be estimated by using random effects models. However, in the present NMA, because of the low number of included trials, the results of the random effects models were unstable, with wide CrIs. For this reason, such models could not be used, and the analyses could not include adjustments for trial-level differences in baseline patient characteristics; hence, the extent to which these factors could have biased the results could not be fully determined. Also, because the random effects models are unstable, the deviance information criterion has limited value for assessing the presence of heterogeneity in treatment effects and to compare model fit between the random effects and fixed effects models.

As with any study that is not randomized, heterogeneity due to unobserved variables could lead to uncontrolled confounding. It should also be noted that, although differences in study design across trials (eg, PALOMA-115 was an open-label trial whereas the study by Mehta et al13 had a crossover element) may have an impact on the magnitude of treatment effects, the NMA methods used in this study cannot control for such differences. Moreover, due to the low number of trials and reporting inconsistencies across trials, the analyses did not control for the type of anticancer therapy received before the diagnosis of mBC.

There were also gaps in the reporting of outcomes or patient characteristics for the HR+/HER2− population in some of the identified studies. For the de novo analysis, the FALCON trial12 (which compared Ful500 vs AI) did not report PFS results for the de novo subgroup; hence, PFS results were extracted for the similar subgroup of patients who were endocrine-therapy naïve with no previous exposure to chemotherapies. The distributions of baseline characteristics were largely similar from a clinical perspective; thus, there is no evidence that the transitivity (ie, similarity of study populations across trials) assumption was violated. However, the low number of trials and the lack of reporting of patient characteristics in some of the trials precluded us from making a strong statistical or clinical argument as to whether differences in baseline characteristics across trials could alter treatment effects. Consistency between direct and indirect comparisons could not be assessed due to the lack of loops in the evidence networks.

Evaluating the presence of publication bias in the included studies is an important component of a meta-analysis, although making judgments about the
The presence of publication bias in an NMA is often difficult and not exempt from subjective considerations. To minimize the risk of publication bias, the search strategy was comprehensive and included several sources of studies, including Medline, Cochrane databases, conference proceedings, and published systematic reviews. The validity of the eligible studies was assessed according to the Cochrane Collaboration’s tool for assessing risk of bias. The risk of bias, based on the domains described in Supplemental Table IV (see the online version at https://doi.org/10.1016/j.clinthera.2018.03.004), was characterized as low to moderate. Other more quantitative approaches to assess risk of bias are unreliable given the small number of studies included in our analyses. Finally, the results were expressed in relative terms (as hazard ratios) rather than in clinical measures such as PFS.

Despite these limitations, the findings of the study represent a synthesis of the available comparative evidence and should help clinicians evaluate different ET options.

CONCLUSIONS
This systematic literature review and NMA provide a comparison of the efficacy of endocrine therapy in monotherapy or combination with TTs as treatments for postmenopausal women with HR+/HER2– mBC, based on evidence available in clinical trials of these treatment regimens. These analyses indicate that women in this population receiving Pal + AI, LEE + AI, FUL + AI, or FUL as first-line treatment had longer PFS than those who received AIs alone. Pal + AI and LEE + AI had the highest probability of being the most effective at delaying progression among all treatments compared in all the patient populations studied herein.

ACKNOWLEDGMENTS
Funding for this research was provided by Novartis Pharmaceuticals Corporation.

Writing support was provided by Cinzia Metallo, PhD, an employee of Analysis Group, Inc. All authors contributed to study design, literature search, data collection, interpretation of the study results, and manuscript writing. Drs. Ayyagari and Patterson-Lomba contributed to data analyses and figure creation.

CONFLICTS OF INTEREST
Drs. Ayyagari, Patterson-Lomba, Xie and Ms. Zhou are employees of Analysis Group, Inc, which has received consultancy fees from Novartis Pharmaceuticals Corporation for this study. Drs. Tang, Dalal and Mr. Chandiwana are employees of Novartis Pharmaceuticals Corporation and own stock or stock options. Dr. Niravath received consulting fees from Novartis. The authors have indicated that they have no other conflicts of interest regarding the content of this article.

SUPPLEMENTARY MATERIAL
Supplemental tables and figures accompanying this article can be found in the online version at https://doi.org/10.1016/j.clinthera.2018.03.004.

REFERENCES


Address correspondence to: Rajeev Ayyagari, PhD, Analysis Group, Inc, 111 Huntington Ave, 14th Floor, Boston, MA 02199. E-mail: rajeev. ayyagari@analysisgroup.com
Supplementary Table 1. Summary of prior treatments.

<table>
<thead>
<tr>
<th>Trial</th>
<th>Mehta 2012</th>
<th>PALOMA-1</th>
<th>PALOMA-2</th>
<th>MONALEESA-2</th>
<th>FALCON</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Anastrozole</td>
<td>Letrozole</td>
<td>Letrozole</td>
<td>Letrozole</td>
<td>Fulvestrant</td>
</tr>
<tr>
<td></td>
<td>fulvestrant</td>
<td>palbociclib</td>
<td>palbociclib</td>
<td>LEE</td>
<td>Letrozole</td>
</tr>
<tr>
<td>Prior therapy, n (%)</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Radiotherapy</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Surgery</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Endocrine therapy</td>
<td>NR</td>
<td>NR</td>
<td>27 (32.0)</td>
<td>28 (35.0)</td>
<td>NR</td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>NR</td>
<td>NR</td>
<td>34 (40.0)</td>
<td>37 (46.0)</td>
<td>NR</td>
</tr>
<tr>
<td>Targeted therapy</td>
<td>NR</td>
<td>NR</td>
<td>213 (48.0)</td>
<td>109 (49.1)</td>
<td>NR</td>
</tr>
<tr>
<td>Immunotherapy</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Other</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Previous hormone therapy, n (%)</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Adjuvant</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Neoadjuvant</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Palliative</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Prevention</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Other</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Prior chemotherapy setting, n (%)</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Adjuvant</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Neoadjuvant</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Palliative</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Other</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
</tbody>
</table>
| Notes:        | *The overall population in the Mehta 2012 trial is HR+, which is broader than HR+/HER2-. HER2- subgroup results were used.

Supplementary Table 2. Pairwise treatment comparison, late progressers analysis. Results show median and 95% CrI of Hazard Ratio (column vs. row).

<table>
<thead>
<tr>
<th></th>
<th>AI</th>
<th>LEE+AI</th>
<th>Pal+AI</th>
</tr>
</thead>
<tbody>
<tr>
<td>AI</td>
<td>1</td>
<td>0.56 (0.43, 0.72)</td>
<td>0.58 (0.46, 0.73)</td>
</tr>
<tr>
<td>LEE+AI</td>
<td>1.80 (1.39, 2.33)</td>
<td>1</td>
<td>1.04 (0.74, 1.48)</td>
</tr>
<tr>
<td>Pal+AI</td>
<td>1.72 (1.38, 2.16)</td>
<td>0.96 (0.68, 1.34)</td>
<td>1</td>
</tr>
</tbody>
</table>

AI: an aromatase inhibitor; LEE: ribociclib; Pal: palbociclib
Supplementary Table 3. Pairwise treatment comparison, de novo analysis. Results show median and 95% CrI of Hazard Ratio (column vs row).

<table>
<thead>
<tr>
<th></th>
<th>AI</th>
<th>LEE+AI</th>
<th>Ful500</th>
<th>Pal+AI</th>
</tr>
</thead>
<tbody>
<tr>
<td>AI</td>
<td>1</td>
<td>0.45 (0.27, 0.76)</td>
<td>0.75 (0.58, 0.97)</td>
<td>0.54 (0.39, 0.74)</td>
</tr>
<tr>
<td>LEE+AI</td>
<td>2.21 (1.32, 3.68)</td>
<td>1</td>
<td>1.66 (0.94, 2.93)</td>
<td>1.19 (0.64, 2.16)</td>
</tr>
<tr>
<td>Ful500</td>
<td>1.33 (1.03, 1.71)</td>
<td>0.60 (0.34, 1.07)</td>
<td>1.16 (0.94, 2.93)</td>
<td>1.19 (0.64, 2.16)</td>
</tr>
<tr>
<td>Pal+AI</td>
<td>1.98 (1.35, 2.58)</td>
<td>0.84 (0.46, 1.55)</td>
<td>1.40 (0.93, 2.12)</td>
<td>1</td>
</tr>
</tbody>
</table>

AI: an aromatase inhibitor; LEE: ribociclib; Pal: palbociclib; Ful500: fulvestrant 500mg

Supplementary Table 4. Risk of bias assessment table.

<table>
<thead>
<tr>
<th>Trial no. (acronym)</th>
<th>Mehta 2012</th>
<th>PALOMA-1</th>
<th>PALOMA-2</th>
<th>MONALEESA-2</th>
<th>FALCON</th>
</tr>
</thead>
<tbody>
<tr>
<td>Was randomisation carried out appropriately?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Not clear</td>
</tr>
<tr>
<td>Was the concealment of treatment allocation adequate?</td>
<td>Not clear</td>
<td>N/A</td>
<td>Not clear</td>
<td>Yes</td>
<td>Not clear</td>
</tr>
<tr>
<td>Were the groups similar at the outset of the study in terms of prognostic factors?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Were the care providers, participants and outcome assessors blind to treatment allocation?</td>
<td>Not clear</td>
<td>N/A</td>
<td>Not clear</td>
<td>Yes</td>
<td>Not clear</td>
</tr>
<tr>
<td>Were there any unexpected imbalances in drop-outs between groups?</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Not clear</td>
</tr>
<tr>
<td>Is there any evidence to suggest that the authors measured more outcomes than they reported?</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Not clear</td>
</tr>
<tr>
<td>Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Not clear</td>
</tr>
</tbody>
</table>
Supplementary Figure 1. Ranking probabilities by treatment, late progressers analysis.

Supplementary Figure 2. Ranking probabilities by treatment, de novo analysis.