Breast Cancer Chemoprevention: A Network Meta-Analysis of Randomized Controlled Trials

Simone Mocellin, Pierluigi Pilati, Marta Briarava, Donato Nitti

Affiliations of authors: Department of Surgery, Oncology and Gastroenterology, University of Padova, Padova, Italy (SM, PP, MB, DN); Istituto Oncologico Veneto, IOV-IRCCS, Padova, Italy (SM); Sant’Antonio Hospital, Padova, Italy (PF).

Correspondence to: Simone Mocellin, MD, PhD, Meta-Analysis Unit, Dept. Surgery Oncology and Gastroenterology, University of Padova, Istituto Oncologico Veneto, IOV-IRCCS, Via Giustiniani 2, 35128 - Padova, Italy (e-mail: simone.mocellin@unipd.it).

Abstract

Background: Several agents have been advocated for breast cancer primary prevention. However, few of them appear effective, the associated severe adverse effects limiting their uptake.

Methods: We performed a comprehensive search for randomized controlled trials (RCTs) reporting on the ability of chemoprevention agents (CPAs) to reduce the incidence of primary breast carcinoma. Using network meta-analysis, we ranked CPAs based simultaneously on efficacy and acceptability (an inverse measure of toxicity). All statistical tests were two-sided.

Results: We found 48 eligible RCTs, enrolling 271,161 women randomly assigned to receive either placebo or one of 21 CPAs. Aromatase inhibitors (anastrozole and exemestane, considered a single CPA class because of the lack of between-study heterogeneity; relative risk [RR] = 0.468, 95% confidence interval [CI] = 0.346 to 0.634), arzoxifene (RR = 0.415, 95% CI = 0.253 to 0.682), lasofoxifene (RR = 0.208, 95% CI = 0.079 to 0.544), raloxifene (RR = 0.572, 95% CI = 0.372 to 0.881), tamoxifen (RR = 0.708, 95% CI = 0.595 to 0.842), and tibolone (RR = 0.317, 95% CI = 0.127 to 0.792) were statistically significantly associated with a therapeutic effect, which was restricted to estrogen receptor–positive tumors of postmenopausal women (except for tamoxifen, which is active also during premenopause). Network meta-analysis ranking showed that the new selective estrogen receptor modulators (SERMs) arzoxifene, lasofoxifene, and raloxifene have the best benefit-risk ratio. Aromatase inhibitors and tamoxifen ranked second and third, respectively.

Conclusions: These results provide physicians and health care regulatory agencies with RCT-based evidence on efficacy and acceptability of currently available breast cancer CPAs; at the same time, we pinpoint how much work still remains to be done before pharmacological primary prevention becomes a routine option to reduce the burden of this disease.

In women, breast carcinoma is the first tumor type by incidence and the second cause of death by cancer (1,2). Surgery remains the only potentially curative approach when the tumor is localized, whereas chemotherapy prolongs the survival of patients with high-risk locally advanced disease or metastatic disease (3,4). Overall, the burden of this disease for the community in terms of mortality, morbidity, psychological stress, and costs is enormous. Therefore, any effort is sorely needed to make breast cancer prevention effective. So far, most results have been obtained with secondary prevention programs, which aim for early detection of the disease (mainly by means of screening mammography), an approach not free of criticisms (5–7).

With regard to primary prevention (ie, all the interventions capable of reducing the occurrence of new cases), several approaches have been proposed and tested, such as lifestyle changes (including smoking cessation and active exercise), prophylactic surgery, and chemoprevention (also known as preventive therapy [8]) (9–11). Although lifestyle changes are broadly recommended for their undeniable effectiveness on a range of malignant and nonmalignant illnesses, the magnitude of their...
impact specifically on breast cancer incidence has been questioned (12). Alternatively, bilateral mastectomy is limited to very-high-risk women (such as those harboring BRCA1/2 germ-line pathologic mutations (13) or those with a personal history of breast cancer (14) and thus cannot have a notable impact on the global incidence of this disease. In the end, the highest expectations reside in chemoprevention (4,8).

In the light of the hormone dependence of the majority of breast carcinomas (at least in the early phase of disease) (15), the most highly researched drug class in breast cancer primary prevention has been that of selective estrogen receptor modulators (SERMs) (16). However, many other potential chemoprevention agents (CPAs) have been investigated, ranging from anti-estrogen drugs (eg, aromatase inhibitors) (17), to micronutrients (eg, vitamins) (18), to compounds investigated for their activity against nonmalignant diseases and then discovered to have potential anticancer effects (eg, anti-osteopenia drugs, anticholesterol drugs, nonsteroidal anti-inflammatory drugs, and antidiabetes drugs) (19,20).

Internationally trusted guidelines (eg American Society of Clinical Oncology [ASCO] [21], National Comprehensive Cancer Network [NCCN] [22], US Preventive Services Task Force [23], National Institute for Health and Care Excellence [NICE] [24]) do recommend the use of some hormone-related CPAs in high-risk women, although the uptake is low mainly because of the associated adverse effects (10,11,25). Randomized controlled trials (RCTs) and meta-analyses of single CPAs support these guidelines (26).

However, to our knowledge, no multiple treatment meta-analysis has yet been published that covers the highest level of evidence (ie, based on the findings of RCTs) on the range of compounds so far tested in this field. Unlike conventional pairwise meta-analysis, the relatively recently implemented “network meta-analysis” (also known as multiple treatment meta-analysis) (27–30) enables investigators not only to combine direct evidence (ie, information deriving from head-to-head comparisons such as treatment A vs treatment B) with indirect evidence (ie, information on the A-vs-B comparison deriving from studies comparing B vs C and A vs C) but also to compare and rank multiple treatments based simultaneously on their efficacy (ability to reduce disease incidence) and acceptability (inversely proportional to CPA toxicity). With the present work we attempted to fill this gap in the medical literature and provide readers with useful information to make evidence-based decisions about the administration of CPAs for the primary prevention of breast cancer in women without prior history of this disease.

Methods

Literature Search

We first performed a systematic review of RCT reporting on the efficacy of any potential CPA to affect the incidence of breast cancer. To this aim we followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (31) and Cochrane guidelines (32). The search consisted of three stages. In the first stage, two authors (SM and PP) independently performed an electronic search of PubMed, Cochrane Collaboration, US Clinicaltrials.gov, EU Clinical Trials Register, and Scopus databases until March 2015. The following medical subject heading terms and text words were used in combination: “breast,” “cancer” or “tumor” or “carcinoma,” “chemoprevention” or “prevention” or “risk” or “incidence” or “preventive therapy,” and “randomized.” In the second stage, the electronic search was enriched with the names of the chemoprevention agents (CPAs) retrieved in the first stage. In the final stage, the references of eligible studies as well as those from review articles and meta-analyses were screened.

According to the so called PICO checklist (32), eligibility criteria were the following: 1) Population: women without breast cancer history; 2) Intervention: any pharmacological compound with hypothesized activity against cancer; 3) Control: placebo (or any CPA in trials comparing different compounds); 4) Outcomes: breast cancer incident cases (beneficial effect, hereafter referred to as “efficacy”) and severe adverse events (toxicity, hereafter referred to as “acceptability”). As regards breast cancer incidence, the primary outcome was any type carcinoma (invasive + non-invasive [in situ carcinoma]). At subgroup analysis, invasive and noninvasive carcinomas as well as estrogen receptor-positive and –negative carcinomas were considered separately. Breast cancer risk of women enrolled in the included trials could not be used as a study-level covariate for network meta-regression analysis because of the low number of available studies. Severe adverse events were defined as treatment-related deaths, life-threatening events (eg, myocardial infarction, stroke, thromboembolism, cancer [eg, endometrial carcinoma]) and any side effect classified as G3-G4.

Only full text articles were included; if multiple publications of the same trial were retrieved, only the most recent and informative publication was included.

Titles and abstracts were independently screened by two reviewers (SM and PP) for potentially relevant studies according to the aforementioned eligibility criteria: After excluding duplicated and nonrelevant studies, the remaining articles were read in full text. Any disagreement was resolved by discussion and consensus with the other two authors (MB and DN). The authors of the original article were contacted when essential data (eg, those necessary to calculate the outcome measures) were unreported.

Risk of Bias Assessment

The Cochrane Risk of Bias Tool was adopted to assess the risk of bias for each RCT (32). This tool includes seven domains: random sequence generation (capturing selection bias), allocation concealment (selection bias), blinding of participants and personnel (performance bias), blinding of outcome assessment (detection bias), incomplete outcome data (attrition bias), selective outcome reporting (reporting bias), and other sources of bias (eg, per-protocol analysis instead of intention-to-treat analysis, sample size calculation, and achievement). Based on the above domains, the included RCTs were classified into one of three categories: low risk, high risk, or unclear risk.

Evidence Grading

The Grades of Recommendation, Assessment, Development and Evaluation (GRADE) system (33) adapted to network meta-analysis (34) was employed to grade the quality of evidence into four levels: high, moderate, low, and very low quality.

The quality can be downgraded by one (serious concern) or two levels (very serious concern) for the following reasons: risk of bias (see above paragraph), inconsistency (unexplained heterogeneity, inconsistency of results), indirectness (indirect population, intervention, control, outcomes), and imprecision (wide predictive intervals, single trial). The quality can also be upgraded by one level because of large sample size (we
chose efficacy relative risk $< 0.5$) and dose-response gradient (e.g., effect increases as the drug dose increases).

### Statistical Analysis

The outcome measure was the relative risk (RR), which is the ratio between the incidence of breast cancer (or severe adverse events) in the experimental arm and that in the control arm. The 95% confidence interval (CI) was calculated as a measure of estimate uncertainty. For direct comparisons, standard pairwise meta-analysis was performed using the DerSimonian-Laird random effects model (35). If a direct comparison was based on two or more studies, result inconsistency across studies (i.e., between-study heterogeneity) was quantified using the $I^2$-squared statistic; heterogeneity was considered low, moderate, or high for $I^2$-squared values under 25%, between 25% and 50% and over 50%, respectively (36).

Heterogeneity assessment also helped us to combine biologically similar CPAs in a single treatment group, which was dictated by the need to minimize the issue of rare categories (i.e., to avoid that a given treatment was represented by one or very few trials). For two or more CPAs to be grouped in a single category, between-study heterogeneity had to be absent.

Random effects network meta-analysis was carried out within a frequentist setting (as an alternative to the Bayesian model) (37,38) to synthesize the available evidence on primary breast cancer prevention agents by combining all the information regarding efficacy and acceptability from different studies. Network meta-analysis combines data from a network of trials testing more than two competing treatments by integrating direct evidence (i.e., deriving from the head-to-head comparison of the treatments of interest, such as A and B) and indirect evidence (i.e., information on the A-vs-B comparison deriving from trials comparing B vs C and A vs C); this increases the estimate precision and produces a relative ranking of all interventions while properly accounting for correlations between effects sizes from multi-arm trials (see also the Methods in the Supplementary Materials, available online) (37,38).

A common heterogeneity parameter ($tau^2$-squared) was assumed across all comparisons. Each summary effect is presented along with its 95% confidence interval (CI) and predictive interval (PrI). The latter (which is calculated using the between-study variance tau-squared) is the interval within which the estimate of a future study is expected to be, thus providing information on the magnitude of heterogeneity.

When combining the results of direct and indirect comparisons, the extent to which these results are inconsistent with each other was examined. In a network of treatments, different direct comparisons may form evidence cycles (also called loops) within which inconsistency is evaluated.

Besides allowing for indirect comparisons, network meta-analysis also provides a ranking probability curve of each treatment (rankogram) by calculating the probability of each arm to achieve the best rank among all. A simple numerical summary to supplement the graphical display of cumulative ranking is to estimate the surface under the cumulative ranking (SUCRA) line for each treatment, which equals one when a treatment is certain to be the best and zero when a treatment is certain to be the worst (39).

Finally, the number needed to treat (NNT; the number of patients to be treated in order to avoid one event of interest) was calculated in order to provide readers with information on the absolute effect of treatment on patients’ outcome. All statistical tests were two-sided. For statistical analysis and graph generation we used Stata 11.2 (StataCorp, College Station, TX). Risk of bias was assessed using Review Manager (RevMan, version 5.1, The Nordic Cochrane Centre: The Cochrane Collaboration, Copenhagen, Norway).

### Results

#### Search Findings

The literature search led to the identification of 48 eligible RCTs (40–77), enrolling 271 161 women randomly assigned to receive either placebo or one of 21 chemically distinct compounds belonging to the following eight pharmacological classes: non-steroidal anti-inflammatory drugs: aspirin; bisphosphonates: alendronate and zoledronate; aromatase inhibitors: anastrozole and exemestane; selective estrogen receptor modulators: tamoxifen, arzoxifene, bazedoxifene, lasofoxifene, and raloxifene; statins: lovastatin, pravastatin, and simvastatin; insulin stimulating agents: metformin and thiazolidinediones (rosiglitazone and pioglitazone); vitamins (vitamin A, vitamin D, vitamin E, and folic acid); synthetic hormones: tibolone (see network plot in Figure 1). The main features of these studies are reported in Table 1.

#### Direct Comparison Meta-Analysis

Results of single RCTs and standard pairwise meta-analysis of direct comparisons are fully reported in the Supplementary Materials (Supplementary Table 1, available online). Out of 21 CPAs, meta-analysis of efficacy was feasible in 11 cases (bazedoxifene, lasofoxifene, raloxifene, tamoxifen, aspirin, pravastatin, simvastatin, vitamin-A, vitamin-D, rosiglitazone, and metformin), for which at least two datasets were available. Among these 11 CPAs, meta-analysis showed that only tamoxifen (summary RR = 0.708, 95% CI = 0.595 to 0.842, $P < .001$) and raloxifene (summary RR = 0.572, 95% CI = 0.372 to 0.881, $P = .01$)

![Figure 1. Network plot: chemoprevention agents tested in randomized controlled trials for their ability to reduce primary breast cancer incidence. All agents were compared with placebo (except for one trial comparing tamoxifen to raloxifene). Circles and connectors are proportional to the number of trials (aromatase inhibitors = 2, arzoxifene = 1, aspirin = 5, bazedoxifene = 2, bisphosphonates = 2, folic acid = 8, lovastatin = 1, lasofoxifene = 2, metformin = 2, pravastatin = 4, raloxifene = 3, simvastatin = 2, tamoxifen = 6, tibolone = 1, thiazolidinediones = 3, vitamin A = 3, vitamin D = 5, vitamin E = 1). ARIN = aromatase inhibitor; BISPH = bisphosphonate; TZD = thiazolidinedione.](https://academic.oup.com/jnci/article-abstract/108/2/djv318/2457808)
## Table 1. Main characteristics of eligible randomized controlled trials

<table>
<thead>
<tr>
<th>Author (reference)</th>
<th>Year</th>
<th>Women: arm 1</th>
<th>Women: arm 2</th>
<th>Breast risk</th>
<th>Menopausal</th>
<th>Median age, y</th>
<th>Setting</th>
<th>Drug regimen</th>
<th>Follow-up, y</th>
<th>Drug regimen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cummings (40)</td>
<td>1998</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Normal</td>
<td>Post</td>
<td>68</td>
<td>Post</td>
<td>800 IU daily</td>
<td>4 years</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Black (41)</td>
<td>2007</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Normal</td>
<td>Post</td>
<td>3093</td>
<td>Post</td>
<td>3.8</td>
<td>55</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Gutzi (42)</td>
<td>2011</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>3788</td>
<td>Post</td>
<td>1.6</td>
<td>37</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Goess (43)</td>
<td>2012</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>3177</td>
<td>Post</td>
<td>5.8</td>
<td>206</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Archer (44)</td>
<td>2009</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>2675</td>
<td>Post</td>
<td>4.5</td>
<td>20</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Arzio (45)</td>
<td>2010</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1985</td>
<td>Post</td>
<td>3.5</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Lкро (46)</td>
<td>2012</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1885</td>
<td>Post</td>
<td>3.5</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Vogel (47)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Decensi (48)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Cook (49)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Tamsulosin (50)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Tamsulosin (51)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Tamsulosin (52)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Tamsulosin (53)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Tamsulosin (54)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Tamsulosin (55)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Tamsulosin (56)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Tamsulosin (57)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Tamsulosin (58)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Tamsulosin (59)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Tamsulosin (60)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Tamsulosin (61)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Tamsulosin (62)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Tamsulosin (63)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Tamsulosin (64)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Tamsulosin (65)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Tamsulosin (66)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Tamsulosin (67)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Tamsulosin (68)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Tamsulosin (69)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Tamsulosin (70)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
<tr>
<td>Tamsulosin (71)</td>
<td>2015</td>
<td>Placebo</td>
<td>Placebo</td>
<td>Mixed</td>
<td>Post</td>
<td>1996</td>
<td>Post</td>
<td>3.7</td>
<td>133</td>
<td>Breast risk</td>
</tr>
</tbody>
</table>

---

**Footnotes:**

were statistically significantly associated with a reduction in breast cancer risk. Subgroup analysis showed that full-dose (20 mg) tamoxifen (summary RR = 0.705, 95% CI = 0.571 to 0.871, P = .001) but not low-dose (5 mg) tamoxifen (summary RR = 0.748, 95% CI = 0.479 to 1.169, P = .20) was effective (Supplementary Table 1, available online). As regards lasofoxifene, the only available RCT showed that only the full dose (0.50 mg) was associated with reduced disease risk (RR = 0.208, 95% CI = 0.079 to 0.544, P = .001).

Considering CPAs with only one dataset available (n = 10), alendronate and zoledronate, and anastrozole and exemestane were grouped together as two CPA classes (bisphosphonates and aromatase inhibitors, respectively) in the light of their homogeneity in terms of both pharmacological features and effect on breast cancer risk (lack of between-study heterogeneity, I-squared = 0%). The meta-analysis performed considering these two CPA groups showed that only aromatase inhibitors were associated with statistically significant risk reduction (summary RR = 0.468, 95% CI = 0.346 to 0.634, P < .001).

Finally, among the remaining CPAs with a single RCT available but without possibility of dataset combination (n = 6), only tibolone (RR = 0.317, 95% CI = 0.127 to 0.792, P = .01) and arzoxifene (RR = 0.415, 95% CI = 0.253 to 0.762, P = .001) were statistically significantly associated with decreased breast cancer incidence.

For the six CPAs (aromatase inhibitors [anastrozole and exemestane, considered a single CPA class], arzoxifene, lasofoxifene, raloxifene, tamoxifen, and tibolone) were statistically significantly associated with a therapeutic effect; standard pairwise meta-analysis was performed also for the acceptability outcome, the findings being reported in Supplementary Table 2 (available online).

### Network Meta-Analysis

In order to compare and rank CPAs based on their efficacy and acceptability, we performed a network meta-analysis of the relevant RCTs. We limited this analysis to the six CPAs (aromatase inhibitors [anastrozole and exemestane, considered a single CPA], arzoxifene, lasofoxifene [0.50 mg], raloxifene, tamoxifen [20 mg], and tibolone), with evidence of efficacy in reducing breast cancer risk based on the above-mentioned direct comparisons data. There were seven mixed comparisons, that is, comparisons exploiting direct and possibly indirect evidence (see contribution matrix in Supplementary Figure 1 [efficacy] and Supplementary Figure 2 [acceptability], available online).

Network meta-analysis results in terms of both efficacy (Figure 2) and acceptability (Figure 3) were highly comparable with those from standard pairwise meta-analysis.

The numbers needed to treat (NNT) were 61, 151, 144, 165, 229, and 174 for aromatase inhibitors, arzoxifene, lasofoxifene, raloxifene, tamoxifen, and tibolone, respectively. These values do not necessarily reflect the magnitude of the relative risk (ie, risk reduction) associated with each CPA, which is because of the varying disease incidence observed in different RCTs (78).

Network meta-analysis generated also 14 indirect comparisons (see contribution matrix in Supplementary Figure 1 [efficacy] and Supplementary Figure 2 [acceptability], available online). As regards efficacy (Figure 2), the only indirect comparison that was statistically significant was that of tamoxifen vs lasofoxifene (RR = 3.39, 95% CI = 1.18 to 9.73, P = .02), suggesting that the therapeutic benefit linked to lasofoxifene is statistically significantly greater than that associated with tamoxifen.
In terms of acceptability (Figure 3), both lasofoxifene (RR = 0.84, 95% CI = 0.74 to 0.96, \( P = .009 \)) and raloxifene (RR = 0.85, 95% CI = 0.76 to 0.95, \( P = .004 \)) were associated with lower incidence of severe adverse effects as compared with aromatase inhibitors; tibolone resulted in more toxicity than aromatase inhibitors (RR = 1.21, 95% CI = 1.02 to 1.43, \( P = .03 \)), arzoxifene (RR = 1.34, 95% CI = 1.13 to 1.59, \( P = .001 \)), lasofoxifene (RR = 1.43, 95% CI = 1.20 to 1.71, \( P < .001 \)), and raloxifene (RR = 1.42, 95% CI = 1.21 to 1.68, \( P < .001 \)); finally, tamoxifen was less acceptable than arzoxifene (RR = 1.18, 95% CI = 1.03 to 1.35, \( P = .02 \)), lasofoxifene (RR = 1.26, 95% CI = 1.10 to 1.45, \( P = .001 \)), and raloxifene (RR = 1.26, 95% CI = 1.11 to 1.42, \( P < .001 \)).

Inconsistency findings were based on the only treatment loop placebo-raloxifene-tamoxifen (see network plot in Figure 1). As regards efficacy, the ratio between relative risks from direct and indirect evidence was not significantly different from one (RR ratio = 1.612, 95% CI = 1.00 to 2.94, \( P = .12 \)).

---

**Table 1: Treatment Effect**

<table>
<thead>
<tr>
<th>Treatment Effect</th>
<th>Mean with 95% CI and 95% PrI</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARINs vs Placebo</td>
<td>1.19 (1.09, 1.29) (1.05, 1.34)</td>
</tr>
<tr>
<td>Arzoxifene vs Placebo</td>
<td>1.07 (0.98, 1.17) (0.94, 1.21)</td>
</tr>
<tr>
<td>Lasofoxifene vs Placebo</td>
<td>1.00 (0.91, 1.10) (0.87, 1.15)</td>
</tr>
<tr>
<td>Raloxifene vs Placebo</td>
<td>1.00 (0.93, 1.08) (0.90, 1.12)</td>
</tr>
<tr>
<td>Tamoxifen vs Placebo</td>
<td>1.26 (1.14, 1.39) (1.09, 1.45)</td>
</tr>
<tr>
<td>Tibolone vs Placebo</td>
<td>1.43 (1.24, 1.65) (1.18, 1.73)</td>
</tr>
<tr>
<td>Arzoxifene vs ARINs</td>
<td>0.90 (0.79, 1.02) (0.76, 1.06)</td>
</tr>
<tr>
<td>Lasofoxifene vs ARINs</td>
<td>0.84 (0.74, 0.96) (0.71, 1.00)</td>
</tr>
<tr>
<td>Raloxifene vs ARINs</td>
<td>0.85 (0.76, 0.95) (0.73, 0.98)</td>
</tr>
<tr>
<td>Tamoxifen vs ARINs</td>
<td>1.06 (0.93, 1.21) (0.89, 1.27)</td>
</tr>
<tr>
<td>Tibolone vs ARINs</td>
<td>1.21 (1.02, 1.43) (0.97, 1.50)</td>
</tr>
<tr>
<td>Lasofoxifene vs Arzoxifene</td>
<td>0.94 (0.82, 1.07) (0.76, 1.12)</td>
</tr>
<tr>
<td>Raloxifene vs Arzoxifene</td>
<td>0.94 (0.84, 1.06) (0.80, 1.10)</td>
</tr>
<tr>
<td>Tamoxifen vs Arzoxifene</td>
<td>1.18 (1.03, 1.39) (0.99, 1.41)</td>
</tr>
<tr>
<td>Tibolone vs Arzoxifene</td>
<td>1.34 (1.13, 1.59) (1.07, 1.67)</td>
</tr>
<tr>
<td>Raloxifene vs Lasofoxifene</td>
<td>1.00 (0.89, 1.13) (0.85, 1.16)</td>
</tr>
<tr>
<td>Tamoxifen vs Lasofoxifene</td>
<td>1.26 (1.01, 1.56) (1.05, 1.52)</td>
</tr>
<tr>
<td>Tibolone vs Lasofoxifene</td>
<td>1.43 (1.20, 1.71) (1.14, 1.80)</td>
</tr>
<tr>
<td>Tamoxifen vs Raloxifene</td>
<td>1.26 (1.11, 1.42) (1.06, 1.49)</td>
</tr>
<tr>
<td>Tibolone vs Raloxifene</td>
<td>1.42 (1.21, 1.68) (1.15, 1.76)</td>
</tr>
<tr>
<td>Tibolone vs Tamoxifen</td>
<td>1.14 (0.95, 1.36) (0.90, 1.43)</td>
</tr>
</tbody>
</table>

![Figure 2](https://example.com/image2.png)  
**Figure 2.** Network meta-analysis results: efficacy of six chemoprevention agents as compared with placebo and comparisons with each other. The treatment effect measure is expressed as relative risk (ratio of breast cancer incidence in the two arms). CI = confidence interval; PrI = predictive interval.

![Figure 3](https://example.com/image3.png)  
**Figure 3.** Network meta-analysis results: acceptability of six chemoprevention agents as compared with placebo and comparisons with each other. The treatment effect measure is expressed as relative risk (ratio of severe adverse effects incidence in the two arms). ARINs = aromatase inhibitors; CI = confidence interval; PrI = predictive interval.
no evidence of inconsistency was found for acceptability data (RR ratio = 1.11, 95% CI = 1.00 to 1.41, P = .40).

Ranking findings for both efficacy and acceptability are reported in Figure 4 (rankograms for the single CPAs are available in Supplementary Figure 3 [efficacy] and Supplementary Figure 4 [acceptability], available online). New-generation selective estrogen receptor modulators (SERMs) arzoxifene, lasofoxifene, and raloxifene appear to be the best candidates for breast cancer prevention therapy because they combine better efficacy with lower incidence of severe adverse events (Figure 4: right upper quadrant of the ranking plot). Aromatase inhibitors and first-generation SERM tamoxifen ranked second and third, respectively. A very low acceptability makes tibolone unlikely to be considered a good option as a breast cancer preventive agent.

At subgroup analysis, when invasive breast cancer only was considered (ie, when in situ carcinoma was excluded) the efficacy findings were similar to those described above for invasive plus noninvasive tumors, as shown in Supplementary Figure 5 (network meta-analysis results, available online) and Supplementary Figure 6 (ranking plot, available online). When only estrogen receptor–positive carcinoma was evaluated, efficacy was even greater for five CPAs (data for tibolone were unavailable) (Supplementary Figure 7, available online), the ranking of which remaining basically unaltered (Supplementary Figure 8, available online). As regards estrogen receptor–negative carcinoma, data were limited to three CPAs (arzoxifene, raloxifene, tamoxifen), with no therapeutic effect either in single RTCs or meta-analysis of relevant trials.

Quality Assessment of Trials and Evidence Grading

None of the RTCs presented evidence for severe risk of bias (see Supplementary Figure 9 [risk of bias by trial] and Supplementary Figure 10 [risk of bias summary], available online). According to the modified GRADE system (see ranking plot in Figure 4), the level of evidence was high for aromatase inhibitors and lasofoxifene. For the latter, despite the availability of only one RCT (which lowers by one level the overall evidence), the summary effect for efficacy was better (ie, lower) than the prespecified 0.5 value (which upgrades by one level the overall evidence). The level of evidence was instead moderate for raloxifene and tamoxifen (the predictive interval of their summary effect crossed the null value) as well as for arzoxifene and tibolone (only one RCT available for each CPA).

Discussion

To our knowledge, we present the findings of the first network meta-analysis in the field of primary prevention of breast cancer by medical treatment. Current national and international guidelines (21–24) are based on the results of single randomized controlled trials (RCTs) as well as standard meta-analyses dedicated to single chemoprevention agents (CPAs) or CPA classes (26). In the present systematic review, we gathered evidence from RCTs assessing the role of 21 CPAs (belonging to eight different pharmacological classes) in more than 270 000 women, which makes the present series the largest ever analyzed in this field.

We ranked the six CPAs for which we found evidence of therapeutic activity (aromatase inhibitors [anastrozole plus exemestane], arzoxifene, lasofoxifene [0.50 mg], raloxifene, tamoxifen [20 mg], and tibolone), by considering simultaneously the two key outcomes: efficacy (ability to lower breast cancer incidence) and acceptability (inversely proportional to the incidence of severe side effects). According to an evidence grading system (GRADE) [33] specifically adapted to network meta-analysis (34), the level of evidence supporting the efficacy of these CPAs was high and moderate for two (ie, aromatase inhibitors and lasofoxifene) and four CPAs (ie, raloxifene, tamoxifen, arzoxifene, and tibolone), respectively. Our results suggest that new-generation selective estrogen receptor modulators (SERMs) arzoxifene, lasofoxifene, and raloxifene are the best candidates for breast cancer prevention therapy (Figure 4).

These three compounds lower breast cancer incidence to a similar (or even higher) degree as compared with other drugs commonly accepted as CPAs for other tumors, such as aspirin for gastrointestinal cancers (79). The ability of these CPAs to (at least) halve the incidence of breast cancer might theoretically lead to better results as compared with secondary prevention strategies such as mammography-based screening (which is associated with a 19% reduction of breast cancer mortality [6]). However, such a comparison is currently unfeasible because chemoprevention studies have thus far focused exclusively on breast cancer incidence, leaving unaddressed the issue of the actual impact of CPAs on disease-related mortality.

Figure 4. Ranking plot representing simultaneously the efficacy (x-axis) and the acceptability (y-axis) of six chemoprevention agents. Bubble size is proportional to the number of randomized controlled trials (reported within the bubble). Bubble color shows the level of evidence. SUCRA = surface under the cumulative ranking (see text for more details).
Despite the wealth of data so far gathered on this subject, several other aspects remain to be elucidated. For instance, tamoxifen is the only CPA with proven activity in premenopausal women: in the light of its lower efficacy and acceptability as compared with other CPAs (Figure 4), the search for CPAs active in premenopausal patients is warranted. Analogously, currently available CPAs basically have no activity against estrogen receptor–negative breast carcinoma, which represents an apparently neglected field of investigation.

In addition, the optimal duration of treatment is unclear. For some CPAs (eg, aspirin [58]), the benefit is more evident after five years; as regards CPAs proven to be active against breast cancer (eg, tamoxifen), findings in the adjuvant setting suggest that 10-year use is more effective than five-year use (80). Unfortunately, no study has addressed this issue in the field of breast cancer primary prevention.

Finally, another key issue is the selection of the candidates who can most benefit from chemoprevention, taking into consideration the side effects associated with CPA administration. Use of risk assessment tools (eg, based on the Gail score and mammographic density) should be greatly encouraged (as is recommended for secondary prevention programs [6]) (81), although currently there is no consensus on the cutoffs to identify women with the highest benefit-risk ratio and thus suitable candidates for chemoprevention. Attempts to exploit molecular profiling (including the assessment of breast cancer–predisposing germ-line polymorphisms) in order to ameliorate the capability of selecting suitable chemoprevention candidates have been made (82,83), according to the principles of personalized medicine; however, none of them has thus far yielded results convincing enough to be implemented in routine clinical practice.

With regard to CPAs for which no evidence of activity against breast cancer is available (eg, aspirin, vitamins, statins), some considerations are needed. For these CPAs, the therapeutic effect has been so far evaluated in unselected women (that is, patients with an average lifetime risk of disease that lowers the likelihood of detecting a substantial benefit). Moreover, the efficacy of some of them (ie, vitamin D) has been recently reanalyzed and found increased in subgroups of patients (ie, after excluding those already taking micronutrient supplements) (84), a fact that calls for a reappraisal of their preventive potential. Finally, CPAs with borderline therapeutic activity (singularly considered) and low (or absent) toxicity might be combined to synergistically increase the magnitude of their preventive benefit and possibly to add to the preventive activity against different illnesses affecting the same person.

Finally, a couple of limitations of this work should be noted. First, despite our efforts to be fully systematic in the literature search, we might have overlooked some relevant articles: The findings from these missing reports might have changed the estimates of the treatment effects we calculated, although our conclusions are mainly based on very robust data (as supported by the narrow confidence intervals of most estimates generated by the network meta-analysis). Second, we could not verify the key assumption of the network meta-analysis, that is, inconsistency between direct and indirect evidence, because virtually all comparisons were between a given CPA and placebo (ie, lack of head-to-head comparisons between treatments): In fact, only one loop (placebo-raloxifene-tamoxifen) was present that showed no inconsistency. Nevertheless, our ranking results might prompt investigators to directly compare different CPAs, which would provide new data for an update of this network meta-analysis.

In conclusion, our findings provide physicians and health care regulatory agencies with RCT-based evidence on efficacy and acceptability of currently available breast cancer CPAs, the use of which still needs to be thoroughly discussed with high-risk women in terms of benefits and risks. At the same time, we pinpointed how much work still remains to be done to make primary prevention a routine option to defeat this disease.

Note
The authors have no conflicts of interest to disclose.

References