Cochrane Library

Cochrane Database of Systematic Reviews

Aromatase inhibitors for treatment of advanced breast cancer in postmenopausal women (Review)

Gibson L, Lawrence D, Dawson C, Bliss J

Gibson L, Lawrence D, Dawson C, Bliss J.
Aromatase inhibitors for treatment of advanced breast cancer in postmenopausal women.
Cochrane Database of Systematic Reviews 2009, Issue 4. Art. No.: CD003370.
DOI: 10.1002/14651858.CD003370.pub3.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Aromatase inhibitors for treatment of advanced breast cancer in postmenopausal women

Lorna Gibson[1], David Lawrence[2], Claire Dawson[3], Judith Bliss[3]

[1]London School of Hygiene and Tropical Medicine, London, UK. [2]Cancer Research UK and UCL Cancer Trials Centre, London, UK. [3]Clinical Trials and Statistics Unit, The Institute of Cancer Research, Sutton, UK

Contact: Judith Bliss, j.bliss@icr.ac.uk.

Editorial group: Cochrane Breast Cancer Group.
Publication status and date: Stable (no update expected for reasons given in 'What's new'), published in Issue 2, 2021.

Citation: Gibson L, Lawrence D, Dawson C, Bliss J. Aromatase inhibitors for treatment of advanced breast cancer in postmenopausal women. *Cochrane Database of Systematic Reviews* 2009, Issue 4. Art. No.: CD003370. DOI: 10.1002/14651858.CD003370.pub3.

ABSTRACT

Background

Endocrine therapy removes the influence of oestrogen on breast cancer cells and so hormonal treatments such as tamoxifen, megestrol acetate and medroxyprogesterone acetate have been in use for many years for advanced breast cancer. Aromatase inhibitors (AIs) inhibit oestrogen synthesis in the peripheral tissues and have a similar tumour-regressing effect to other endocrine treatments. Aminoglutethimide was the first AI in clinical use and now the third generation AIs, anastrozole, exemestane and letrozole, are in current use. Randomised trial evidence on response rates and side effects of these drugs is still limited.

Objectives

To compare AIs to other endocrine therapy in the treatment of advanced breast cancer in postmenopausal women.

Search methods

For this update, the Cochrane Breast Cancer Group Specialised Register and the Cochrane Central Register of Controlled Trials (CENTRAL) and relevant conference proceedings were searched (to 30 June 2008).

Selection criteria

Randomised controlled trials in postmenopausal women comparing the effects of any AI versus other endocrine therapy, no endocrine therapy, or a different AI in the treatment of advanced (metastatic) breast cancer. Non-English language publications, comparisons of the same AI at different doses, AIs used as neoadjuvant treatment, or outcomes not related to tumour response were excluded.

Data collection and analysis

Data from published trials were extracted independently by two review authors and cross-checked by a third. Hazard ratios (HR) were derived for analysis of time-to-event outcomes (overall and progression-free survival). Odds ratios (OR) were derived for objective response, clinical benefit, and toxicity.

Main results

Thirty-seven trials were identified, 31 of which were included in the main analysis of any AI versus any other treatment (11,403 women). No trials were excluded due to inadequate allocation concealment. The pooled estimate showed a significant survival benefit for treatment with an AI over other endocrine therapies (HR 0.90, 95% CI 0.84 to 0.97). A subgroup analysis of the three commonly prescribed AIs (anastrozole, exemestane, letrozole) also showed a similar survival benefit (HR 0.88, 95% CI 0.80 to 0.96). There were very limited data to compare one AI with a different AI, but these suggested an advantage for letrozole over anastrozole.

AIs have a different toxicity profile to other endocrine therapies. For those currently prescribed, and for all AIs combined, they had similar levels of hot flushes and arthralgia; increased risks of rash, nausea, diarrhoea and vomiting; but a 71% decreased risk of vaginal bleeding and 47% decrease in thromboembolic events compared with other endocrine therapies.

Authors' conclusions

In women with advanced (metastatic) breast cancer, aromatase inhibitors including those in current clinical use show a survival benefit when compared to other endocrine therapy.

PLAIN LANGUAGE SUMMARY

Aromatase inhibitors for treatment of advanced breast cancer in postmenopausal women

Advanced (or metastatic) breast cancer is cancer that has spread beyond the breast and regional lymph node areas. Breast cancer can progress to metastatic disease despite the person undergoing a range of therapies given after initial treatment, such as surgery, chemotherapy or radiation therapy. Metastatic breast cancer is treatable but it is not curable. Most breast cancer is sensitive to the female hormone oestrogen. Sensitive cancer cells need oestrogen to stay alive and removal of oestrogen from the body, or stopping any circulating oestrogen getting to the cancer cells, is very effective treatment for hormone-sensitive breast cancers. Endocrine (hormonal) therapy removes the influence of oestrogen on breast cancer cells. Hormonal treatments for advanced breast cancer include tamoxifen, the progestins megestrol acetate and medroxyprogesterone acetate, and aromatase inhibitors (AIs). AIs reduce the body's ability to make (synthesise) oestrogen and have tumour-regressing effects. The AIs in current clinical use include anastrozole, exemestane, and letrozole.

The aim of this systematic review was to compare AIs to other endocrine therapy in the treatment of advanced (metastatic) breast cancer. A systematic search was conducted which identified 37 controlled trials in which over 14,000 women were randomised to treatment groups. Treatment with an AI improved survival for women with metastatic disease by 10%. The overall benefits on disease-free survival and response of the tumour were however unclear based on the studies included in this review. Trials using AIs as first-line and second-line therapy reported benefits of therapy that varied with the different AIs and measures of effectiveness. We were unable to identify specific subgroups of women who may benefit from AI use.

Toxicity (negative side effects) was not well reported in the trials. Where it was reported, there was variation as to the method used for reporting, the type of toxicities reported, as well as the criteria used to assess toxicity. Nevertheless, toxicity data were available for 26 of the 32 trials where an AI was compared with a non-AI. AIs had similar levels of arthritic pain (arthralgia) and hot flushes (especially when compared to tamoxifen); increased risks of rash, diarrhoea, nausea and vomiting; but decreased risk of vaginal bleeding and blood clotting (thromboembolic) events compared with other endocrine therapies. Limited quality of life (QOL) data were provided and, as such, no conclusions can be drawn by this review as to the effect on QOL related to an AI versus a non-AI. This is due to the differences between participants and the side-effect profiles of the agents used, different methods of drug application (injection versus tablets), and use of four different QOL instruments at several different timepoints, some which provided results of responders versus non-responders rather than by treatment group. Some QOL measures were based on clinician-reported rather than patient-reported symptoms.

BACKGROUND

Description of the condition

Breast cancer is the most common cause of cancer and cancer mortality in women worldwide (Ferlay 2000). Metastatic breast cancer occurs when the cancer has spread beyond the breast and regional node areas. Breast cancer can progress to metastatic disease despite a range of adjuvant systemic therapies. Once breast cancer is metastatic it is no longer curable, but it is treatable. The aim of any further treatment is to improve the individual's quality and length of life.

Description of the intervention

Endocrine (hormonal) therapy removes the influence of oestrogen on breast cancer cells, preventing the cancer cells from growing and spreading. It has been shown to improve survival in early breast cancer EBCTCG 2005. Hormone dependency of breast cancer was first demonstrated in the 19th century by a Glasgow surgeon, Thomas Beatson, who achieved temporary regression of metastatic disease by oophorectomy (Beatson 1896). Other early methods of therapy consisted of adrenalectomy and hypophysectomy. These procedures have largely been superseded by effective hormonal treatments. Most endocrine therapies either block the binding of oestrogen to its receptor, for example tamoxifen, or reduce serum and tumour concentrations of oestrogen, for example aromatase inhibitors (AIs). A positive initial response to endocrine treatment is a good indication for use of second and even third-line endocrine therapy, until the disease becomes hormone resistant (Roseman 1997). The most important predictor of response to hormone therapy is the oestrogen receptor (ER) status of the original tumour.

How the intervention might work

Currently, the most widely-used endocrine therapy for treatment of hormone-sensitive metastatic disease is tamoxifen (Howell 1997). Tamoxifen is an oral, non-steroidal competitive ER antagonist. Tamoxifen, however, also has an agonist effect and although patients may relapse and develop acquired resistance to tamoxifen, this does not mean that they will not respond to other endocrine therapy.

Other endocrine therapies used in this setting are fulvestrant, megestrol acetate (MA), and medroxyprogesterone acetate (MPA). Fulvestrant is an ER antagonist that both downregulates and degrades the ER and reduces progesterone receptor content but, unlike tamoxifen, does not have an agonist effect. It can be used as a treatment for tamoxifen-resistant advanced disease or after failure of treatment with an AI, so is an alternative second choice to an AI. MPA and MA are oral progestogens which have been shown to have significant antitumour activity after failure of other endocrine therapies in postmenopausal patients.

In postmenopausal women, oestrogen is no longer produced in the ovaries but androgens (mainly from the adrenal glands) are converted to oestrogens in peripheral tissue by the enzyme aromatase (Miller 1996a). Aromatase inhibitors (AIs) are a class of compounds that act systemically to inhibit oestrogen synthesis in tissues. AIs are of two types, reversible and irreversible; both types of inhibitors compete with normal substrates for binding on the enzyme. The non-competitive inhibitors (which are steroidal) leave the enzyme permanently inactivated (Ibrahim 1995).

AIs are classified as either first, second or third generation (Table 1). Aminoglutethimide (AG) was the first AI and although effective it was poorly tolerated. This was supplanted by 4-hydroxy androstenedione (formestane), which was better tolerated. Third generation AIs fall into two principal categories of (a) non-steroidal, reversible triazole derivatives (anastrozole, fadrozole, letrozole, vorozole) and (b) steroidal, irreversible inhibitors (exemestane). The most widely used AIs are currently anastrozole, exemestane, and letrozole.

Why it is important to do this review

AIs have a different toxicity profile to other endocrine therapies, although some side effects that mimic menopausal symptoms due to depletion of oestrogen are the same, such as hot flushes and sweating. Adverse events particular to AIs include stomach upsets (nausea, vomiting, diarrhoea), rash, and arthralgia. In particular, AG is poorly tolerated and can cause drowsiness, fever, and inhibition of cortisol synthesis. Formestane, although generally well tolerated as a treatment, results in a local reaction around the injection site. Of the other endocrine therapies, tamoxifen can cause endometrial changes including vaginal bleeding and increased risk of thromboembolic events. Side effects with progestogens are usually mild but may include hot flushes, night sweats, nausea and indigestion, fluid retention, weight gain, and headaches as well as an increased risk of thromboembolism. Fulvestrant can have similar oestrogen deprivation side effects, injection site reactions, vomiting and diarrhoea.

AIs are now increasingly being used in the treatment of early breast cancer, which may have an impact on their use in advanced (metastatic) disease.

OBJECTIVES

This systematic review aimed to compare AIs to other endocrine therapy in the treatment of advanced (metastatic) breast cancer in postmenopausal women.

METHODS

Criteria for considering studies for this review

Types of studies

Only randomised controlled trials in the following populations were included.

Types of participants

a) Included:

- postmenopausal women with advanced (stage 3) or metastatic (stage 4) breast cancer either at diagnosis or upon relapse;
- oestrogen receptor (ER) positive or status unknown.

b) Excluded:

- local recurrence only;
- with no restrictions on metastatic site or age of the women;
- inclusion not limited to use of an AI as first-line therapy.

Types of interventions

- Any AI versus any other endocrine treatment

- Any AI versus no endocrine treatment
- Any AI plus other endocrine treatment versus other endocrine treatment alone
- Direct comparison between different AIs

Types of outcome measures

Outcome measures were defined a priori as follows.

Primary outcomes

Overall survival (defined as time from date of randomisation to date of death from any cause)

Secondary outcomes

1. Progression-free survival (defined as time from date of randomisation to disease progression), also known as time to progression
2. Clinical response rate, comprising objective response (those women with either complete or partial shrinkage of the tumour) and clinical benefit (objective response plus stable disease for more than 24 weeks)*
3. Treatment toxicity (particularly AI related)
4. Quality of life (QOL), where available and comparable
5. Dropout rate
6. Time to treatment end (stopped or changed due to toxicity)

* International Union Against Cancer (UICC) guidelines were used for evaluation of these criteria (Hayward 1977).

Subgroup analyses

The following subgroup analyses were considered:

- first-line therapy (where the AI was given as initial therapy for advanced disease);
- second-line therapy (where the advanced disease had already been treated with a different AI or another endocrine therapy);
- ER positive versus ER unknown;
- according to site of distant metastases and differential treatment effect.

Search methods for identification of studies

Only English language publications were included.

Electronic searches

1. The Cochrane Breast Cancer Group Specialised Register.

For the first published version of this review (Gibson 2007), the Cochrane Breast Cancer Group Specialised Register was searched (December 2004, 30 September 2005). For this update, additional searches were conducted (30 June 2008). Details of the search strategy used by the group for the identification of studies and the procedure used to code references are outlined in the group's module (www.mrw.interscience.wiley.com/cochrane/clabout/articles/BREASTCA/frame.html). Studies coded as "advanced" and "endocrine therapy" were extracted for consideration.

2. The Cochrane Central Register of Controlled Trials (CENTRAL) (*The Cochrane Library* 2008, Issue 2). See Appendix 1.

Searching other resources

Reference sections of each published paper were searched for additional studies. Conference proceedings were also searched but abstracts, apart from one, were not included. The one that was included (Schmid 2001) had no corresponding publication but there was adequate information in the abstract for the trial to be included.

Data collection and analysis

Selection of studies

Two authors (CLD, LJG) assessed all trials identified through the search strategy and independently decided on eligibility; any differences were resolved by discussion and confirmed by a third author (DJL). Final confirmation of inclusion was made by two authors (LJG, DJL). Any exclusions have been justified and documented in the table Characteristics of excluded studies.

Data from unpublished trials are not included in the review but these are included in the table of ongoing trials. For these, information was obtained from the trial protocol or other available source. Authors were approached for missing or additional information however only two replies were received out of six contacted.

Data extraction and management

Two review authors (CLD, LJG) extracted data independently using data extraction forms designed for this purpose. Data extracted included details of treatment arms and patient numbers, baseline patient characteristics, tumour response rates, time to progression, median survival, and median follow up. Data on toxicity and quality of life were extracted at a later date. The authors were not blinded to the source of the document for article selection or data extraction. A third author (DJL) assessed the data collected to ensure consistency and accuracy. Any differences were resolved by discussion. Data on study quality were extracted as described in 'Assessment of the methodological quality'. Hazard ratios and their associated variances were extracted for all measures available. If a hazard ratio and confidence interval were not reported, these values were calculated (Parmar 1995). Of the report authors (n = 8) who were contacted for supplementary information on the primary endpoints, only two replied (and the data were not available). For the updated review, data extraction was performed independently by two of the review authors (CLD, DJL).

Assessment of risk of bias in included studies

Two review authors (LJG, CLD) independently assessed the quality of all trials deemed eligible and discrepancies were resolved by discussion. The quality of each trial was assessed based on reports in the publication regarding:

- quality of allocation concealment;
- comparability between the baseline characteristics of the treatment arms;
- inclusion of all randomised participants in the analysis;
- details of dropouts.

Randomisation was assessed by grading the allocation concealment (for example blinded, stratified) as: A = adequate, B = unclear, C = inadequate (see Characteristics of included studies). It was not possible to accurately assess the quality of randomisation

in all trials due to lack of information in the published articles. Any imbalance between treatment arms, both in numbers and characteristics, was taken into account in the grading.

Intention-to-treat statements: analyses that were stated to be by intention to treat and included all randomised patients for the primary endpoint. However, it is common practice to report response variables, that is clinical benefit and objective response, only on 'assessable' patients. We have reported these outcomes on both assessable and randomised patients.

Description of the eligibility and exclusion criteria: all trials described in detail the patient characteristics of those patients eligible for the trial. The table 'Characteristics of included studies' includes information on the balance of baseline characteristics, details of patients excluded after randomisation, definitions of the outcome measures, duration of follow up, and median length of follow up.

Measures of treatment effect

The most complete dataset feasible was assembled. Data were, however, only available for the following endpoints: overall survival, progression-free survival, clinical benefit, objective response, and toxicity.

Overall and progression-free survival were analysed using time-to-event methods and for this the hazard ratio (HR) is the most appropriate statistic. If a HR and corresponding confidence intervals (CI) were not reported, these values were calculated indirectly using median time to event (progression or survival) and the number of events extracted from the published Kaplan-Meier curves, following the method of Parmar 1995. A weighted average of survival duration across trials was then calculated. Ratios of treatment effects for time to event were reported so that HRs less than 1.0 favoured the AI regimen.

Response rates were obtained from the tables of best response presented for each trial. Response has been analysed based on assessable (not randomised) patients as most of the trials included in this review only reported responses in this way. As a sensitivity analysis, we also analysed results by intention to treat (ITT); there was no difference. Response rates were analysed as dichotomous variables (for example objective response compared complete or partial response versus stable disease or no response). An odds ratio (OR) and its associated 95% CI were calculated for each trial and a pooled OR derived. Ratios of treatment effects on response were reported so that ORs less than 1.0 favoured the AI regimen. In this case, the 'event' is in effect 'not getting an objective reponse' or 'not getting a clinical benefit'.

Results are presented graphically and all figures follow the same format. Each trial is presented as a single line within each category. The point estimate of the treatment effect is represented by a square, the size of which is proportional to the size of the trial. The associated 95% CI is included as a horizontal line. The summary in each category is represented by a diamond, the north-south axis is the pooled estimate and the east-west axis is the 95% CI.

Unit of analysis issues

There were no unit of analysis issues in this review.

Dealing with missing data

The number of actual dropouts was very difficult to quantify as the quantity and quality of reporting varied greatly. Only 10 trials gave full details by treatment arm. Three trials quoted the number of patients withdrawn due to toxicity as "a small number" (Buzdar 1996b; Buzdar 1996c; Kaufmann 2000). Thus the patients that could be confidently identified as lost to follow up, refusals, or withdrawals totalled 62.

Toxicity

Not all toxicities (also known as side effects or adverse events) were reported in this review. We selected eight predefined toxicities from expert experience, reflecting side effects specific to AIs (nausea, vomiting, diarrhoea, rash, arthralgia) and other hormonal treatments (hot flushes, vaginal bleeding, thromboembolic events). Each side effect was analysed as a dichotomous variable (yes or no) with the effect of the AI considered separately to that of the comparator. This was deemed the most informative method of presentation as the different comparators have different toxicity profiles, whereas AIs have similar toxicity profiles. An OR and its associated 95% CI were calculated for each trial and a pooled OR derived. Ratios of treatment effects for toxicity were reported so that ORs less than 1.0 favoured the AI regimen.

Assessment of heterogeneity

Statistical heterogeneity between trials was assessed using the Chi^2 statistic. However, there were cases where the value for the I^2 statistic was high but the Chi^2 test was not statistically significant; we advise caution in interpreting these results. Evidence of heterogeneity between trials was identified for tumour response rates and progression-free survival though not overall survival, which seems less susceptible to heterogeneity. The reasons for this are unknown but this statistical heterogeneity may be explained by clinical heterogeneity. It is possible that outcomes involving subjective endpoints, that is tumour response, may be subject to variation whereas the hard endpoint used in the survival analysis is unequivocal. With progression-free survival, the trials were undertaken in populations that varied considerably. For example, some trials were using the AI as first-line treatment, some as second-line treatment, and in other trials as mixed first- and second-line. Other contributory factors may be the difference in dosage of some AIs and significant differences in the proportion of patients who were truly hormone receptor positive. We stress that as this review describes a very mixed range of studies of mixed patient populations, carried out over 30 years, the relative effect between treatment arms would still be consistent even with this mix of different patient groups.

Assessment of reporting biases

There is a lack of reporting of overall survival information compared to tumour response. Many of the trials were carried out over 10 years ago but there have been no subsequent publications with updated (or any) survival information.

Data synthesis

The Cochrane Review Manager Software (RevMan5) was used to analyse the data.

A Mantel-Haenszel fixed-effect model was used for the primary analyses (see the Cochrane Handbook for Systematic Reviews of

Interventions) unless there was significant heterogeneity, in which case a random effects model was used (Higgins 2003). A fixed-effect model was used for all overall survival analyses, and all analyses for any AI versus a different AI. For any AI versus a non-AI, and current AIs versus non-AIs, a random-effects model was used for progression-free survival, clinical benefit, and objective response. When an AI was used as first-line therapy, a random-effects model was used for the clinical response variables but not for progression-free survival; whereas a fixed-effect model was used for the clinical response variables for second-line therapy.

A pooled analysis was performed in each group, but the results from each AI were considered separately within the same group, where possible. Trials were pooled by type of AI for survival, progression-free survival, clinical benefit, and objective response outcomes. For toxicity, the data were pooled by type of comparator, that is tamoxifen, MA, MPA, or fulvestrant, as the toxicities of different AIs are similar due to their mode of action. This approach was considered to be more informative due to differences between the AIs (first versus second or third generation; steroidal versus non-steroidal). Post hoc, it was also decided to separately present the pooled results for the AIs that are in current clinical use (by definition the newer, third generation AIs) as this is more relevant to the clinical situation today. The AIs included were: aminoglutethimide (first generation); formestane (second generation); and anastrozole, exemestane, fadrozole, letrozole and vorozole (third generation). The non-AIs included were: megestrol acetate (MA), tamoxifen, fulvestrant, medroxyprogesterone acetate (MPA), and hydrocortisone (HC).

Subgroup analysis and investigation of heterogeneity

In all cases, tests for heterogeneity have been performed across all trials and in each of the treatment groupings outlined above. Some of the trials that were pooled used different doses of AI, which may have contributed to some of the heterogeneity. Instances of statistically significant heterogeneity are discussed in the results section.

Sensitivity analysis

All analyses were based on the intention-to-treat (ITT) principle as far as was possible, that is comparing all women allocated to one treatment versus all those allocated to the other irrespective of compliance. Thus the results may slightly underestimate any treatment effects. However, analysis of response used the number of assessable women as the denominator as this is the accepted method. As a sensitivity analysis, both denominators were used (see Figures) and there was no major difference for response when comparing assessable to ITT. For statistical tests a P value of less than 0.05 was considered to denote statistical significance.

RESULTS

Description of studies

Some references were excluded because they were either non-English language papers, reviews, non-randomised studies, or conference proceedings without the addition of published data. The exception to this was the conference abstract by Schmid 2001 which was included as it presented several of the endpoints relevant to this review in abstract form; there is no published paper for this trial. Where a trial compared two doses of an AI with a comparator, the trial was included using the arm with the standard or most commonly used dose of that particular AI versus the comparator. For anastrozole this was 1 mg; and for fadrozole it was 2.5 mg, or 2 mg if 2.5 mg was not used.

Results of the search

The original search (Gibson 2007) yielded 152 English language references, of which 133 were possibly eligible. Twenty-five of the 133 references, relating to 22 trials, were excluded as they compared the same AI at different doses. Fifty relevant references were identified relating to 25 randomised trials which fulfilled the eligibility criteria. An additional five references for five trials were identified by the authors from reference lists in papers and reviews. The updated search (June 2008) yielded a further 54 references of which 17 were possible inclusions. This resulted in a further seven trials being assessed as eligible for inclusion.

Included studies

We included 37 trials which randomised 14,060 women. There was a great deal of variation across the trials. Trials ranged in size from 60 (Kleeberg 1997) to 1021 women (Bonneterre 2001). Fourteen trials randomised patients from multiple countries; of the remaining 23 trials, three were limited to the UK, two were from Spain, two from South Africa, two from Switzerland, six from the US, and one each from North America (US and Canada), Denmark, France, Germany, Greece, Japan, Norway, and Switzerland. The country was not formally reported in 14 trials but surmised from the addresses of the authors.

In 32 trials comparing AIs with non-AIs,11,710 women were randomised; 2350 women were randomised in five trials of one AI versus a different AI. It should be noted that seven trials included two different doses of an AI compared with a third comparison. Data for 12,883 women were included in this review.

Of the 32 trials comparing AIs with non-AIs, 11 used the first generation AI aminogluthetimide, three used the second generation AI formestane, and 18 used a third generation AI (anastrozole: four trials; exemestane: three trials; fadrozole: six trials; letrozole: four trials; vorozole: one trial). In these, tamoxifen was the comparator in 12 trials, MA in 13 trials, MPA in four trials, hydrocortisone (HC) in one trial, and fulvestrant in two trials.

The five trials of AIs versus a different AI compared letrozole versus anastrozole, aminoglutethimide, atamestane, or fadrozole; and anastrozole versus formestane.

The AI arm in some of the older trials (Alonso-Munoz 1988; Canney 1988; Ingle 1986; Powles 1984; Rose 1986; Russell 1997) did not compare an AI by itself but in combination with another treatment. One very recent trial (Goss 2007) compared an AI versus a new AI (atamestane) in conjunction with a selective oestrogen receptor modulator (SERM).

In 11 of the 37 trials (randomising 3876 women) in which any AI was used as first-line treatment versus any other comparator, tamoxifen was compared in all trials. In 19 of 37 trials (7413 randomised women) any AI was compared with any comparator as second-line therapy. In four trials the AIs were used as both first- and second-line treatments within the trials, but as the data were not split by this variable they were not included in these comparisons.

Data for all endpoints were not available in the published reports. Where data were unavailable, authors were approached for supplementary data. Five principal endpoints with sufficient data were identified: overall survival, progression-free survival, response (either based on clinical benefit or objective response), and treatment toxicity. Likewise, data were not available in the published reports for all subgroups proposed in the review protocol. The AI versus any non-AI comparison had enough data for all five endpoints as well as a subgroup consisting of data from the three most commonly prescribed AIs (anastrozole, exemestane, and letrozole). In addition, the results of four of the five endpoints (not toxicity) outlined above are presented in three separate groups based on: individual AIs versus different AIs, AIs used as first-line treatment only, AIs used as second-line therapy only.

Time to treatment end

No trial had a fixed treatment period. However, all but two of the trials (Leitzel 1995; Samonis 1994) reported on at least one of the following: time to progression, time to failure or time to death, or both of the latter.

Excluded studies

Non-randomised studies, trials in premenopausal women, and non-English language publications were criteria for exclusion from the review. Trials which compared two different doses of the same AI were also excluded (see the table 'Characteristics of excluded studies').

Risk of bias in included studies

Thirty-seven randomised trials were included in this review. One of the included trials did not have data on the primary or secondary endpoints so could not be included in any analysis (Leitzel 1995). It should be noted that trials by the author of one of the included trials (Bezwoda 1998), relating to high dose chemotherapy, have been found to include falsified data. However, no such findings have been reported for trials included in this review. There was therefore no reason to exclude the trial. Analysis was performed with and without this trial and there was no difference in the pooled results, although for clinical benefit the result became just significant.

It was not possible to accurately assess the quality of all trials due to lack of information in the published articles. Allocation concealment was rated as adequate in 23 trials but there were insufficient details of the allocation concealment in the remaining 14 and so they were labelled as unclear. Of these, no randomisation method was given in eight trials and four were reported to have parallel groups. No trials were deemed to have inadequate allocation concealment, from the information given in the published papers, and none were excluded for this reason. Six trials were double-blind, double-dummy; seven were double-blind; one was double-blind in one arm but open in the other (Buzdar 1996a); and one (which consisted of two trials analysed together) was double-blind in one and open in the other trial (Mauriac 2003).

Baseline characteristics were not commented upon in 12 trials, five trials commented on a slight imbalance. One trial (Buzdar 1996a) had an imbalance in the treatment arm but this was believed to be an artefact. Another trial (Lundgren 1989) reported that "the two groups were well balanced with regard to the most important prognostic variables, with the exception of main metastatic site". All other trials reported balanced baseline characteristics in all arms.

Summary of numbers of women used in the analysis

Women randomised, all arms: 14,060
Women randomised, included arms: 12,883
Women randomised, assessable (for response): 11,111

Effects of interventions

Over 12,000 women were randomised to the included arms of 37 trials but time-to-event data were only available for about half of them. The results of the meta-analysis should be interpreted bearing this in mind.

Aromatase inhibitors (AIs) versus any non-aromatase inhibitor
Of the 32 trials comparing an AI versus a non-AI, one had no data on response or survival by treatment arm although these were included as endpoints (Leitzel 1995). Of the remaining 31 trials, data were available on overall tumour response rates in all 31, clinical benefit in 26, progression-free survival in 11 and overall survival in 13 trials. For overall survival, the reported figures were available from the publications for six trials (Bonneterre 2001; Buzdar 1996a; Buzdar 2001; Ingle 1986; Dombernowsky 1998; Thuerlimann 1996) and were calculated for seven trials (Bezwoda 1998; Gale 1994; Goss 1999; Kaufmann 2000; Milla-Santos 2003; Rose 1986; Russell 1997). In terms of progression-free survival, HRs were reported in the publications of five trials (Bonneterre 2001; Buzdar 2001; Chia 2008; Ingle 1986; Mourisden 2001). The remaining six trials (Dombernowsky 1998; Goss 1999; Kaufmann 2000; Mauriac 2003; Russell 1997; Thuerlimann 1997) had sufficient data for calculation of the HRs.

1. Overall survival Data on survival were available in 13 trials reporting an estimated 2776 events in 4789 women. No data were available for formestane. The pooled HR of 0.90 (95% CI 0.84 to 0.97) showed a statistically significant 10% benefit of treatment (P = 0.007) with an AI, with a consistent effect across all subgroups. Data on individual AIs were sparse and no conclusions could be drawn.

2. Progression-free survival
Data on progression-free survival (PFS) were available in 11 trials reporting an estimated 4391 events in 5890 women. PFS was not statistically significantly associated with the use of an AI (HR 0.98, 95% CI 0.84 to 1.13). This overall effect is virtually uninterpretable due to the significant heterogeneity (P < 0.00001) by type of AI and also within specific AIs.

3. Proportion of women with clinical benefit (8789 assessable women)
Data were available for seven AIs (aminoglutethimide, formestane, anastrozole, exemestane, fadrozole, letrozole, vorozole) from 27 trials. Approximately one third of the data came from three trials (Bonneterre 2001; Mauriac 2003; Mourisden 2001). The AIs were shown to be superior to the non-AIs (OR 0.87, 94% CI 0.77 to 0.99) and there was statistically significant heterogeneity (P = 0.008) across trials.

4. Proportion of women with an objective response (9595 assessable women)
Thirty-one trials reported objective response. Data were available for seven AIs (aminoglutethimide, formestane, anastrozole, exemestane, fadrozole, letrozole, vorozole). The pooled OR suggested no statistically significant effect of treatment with an AI (OR 0.88, 95% CI 0.77 to 1.01) and again there was statistically significant heterogeneity (P = 0.02). Of the individual AIs, only

letrozole was associated with a statistically significant benefit over the non-AI (OR 0.65, 95% CI 0.51 to 0.82) in the 1637 women randomised (Buzdar 2001; Dombernowsky 1998; Mourisden 2001; Schmid 2001).

5. Toxicity

Not all trials provided data on toxicity and there were inconsistencies among trials where it was reported. Toxicity data were available for only 26 of the trials comparing an AI with a non-AI. Within trials, the reported toxicities varied both in the number or range and types of toxicities reported as well as the criteria used for reporting. Some trials reported predefined or selected toxicities (Bonneterre 2001; Kaufmann 2000; Mauriac 2003), some chose to report toxicities occurring in a certain minimum percentage of participants (Bezwoda 1998; Buzdar 2001; Chia 2008; Dombernowsky 1998; Goss 1999; Mauriac 2003; Mourisden 2001), some used worst toxicity grades (Falkson 1996; Thuerlimann 1996; Thuerlimann 1997) or major toxicity (Canney 1988); one reported toxicity grades 1 to 4 separately (Paridaens 2003), one used common toxicities (Buzdar 1996a) though what this meant was not defined, two reported adverse experiences (Buzdar 1996b; Buzdar 1996c), and two reported all toxicities (Freue 2000; Rose 1986). Eight trials did not state which reporting criteria they used. In addition, one trial (Perez Carrion 1994) only reported on the toxicities considered to be treatment related and has not been included. For the trial of an AI against fulvestrant (Mauriac 2003), data on toxicity were obtained from different sources. The combined analysis of the two trials 0020 and 0021 reported predefined events and data on hot flushes and thromboembolic events were available. The separate publications of the results of 0020 and 0021 detailed toxicities occurring in 10% or more of the participants. Trial 0020 reported data on both nausea and vomiting so these were combined with these data from 0021. In addition, trial 0021 had data on the frequency of diarrhoea and rash.

Despite the different reporting criteria the data were pooled. This must be borne in mind when looking at the absolute numbers. The analyses are reported according to the comparator due to the different toxicity profiles of each. An overall pooled result for AI versus non AI is not provided.

Hot flushes

Hot flushes were the specific toxicity that was most widely reported. Data on hot flushes were available from 20 of the 32 trials with 8306 women. Of these, seven compared an AI with tamoxifen, 10 with MA, two with fulvestrant, and one with MPA. The use of an AI had a very similar risk of hot flushes to tamoxifen and fulvestrant. The AI was associated with statistically signficantly more reports of hot flushes than with MA (OR 1.73, 95% CI 1.40 to 2.14) but less than with MPA (OR 0.20, 95% CI 0.06 to 0.73), which had data from only one very small trial.

Nausea

Data on nausea were available from 18 trials with 7895 women. Another two trials reported data on nausea and vomiting combined. Of the 18, six compared an AI with tamoxifen, nine with MA, two with fulvestrant, and one with MPA. AIs were associated with a statistically significant increase in risk of nausea compared to MA (OR 1.77, 95% CI 1.33 to 2.35) but there was no statistically significant difference between AIs and tamoxifen (P = 0.32) or fulvestrant (P = 0.96).

Vomiting Two trials had data on nausea and vomiting combined and so were not included. Data on vomiting were available from two trials comparing AIs with tamoxifen, five versus MA, and one versus fulvestrant for a total of 4404 women. The AI was statistically significantly worse when compared to MA (OR 2.03, 95% CI 1.42 to 2.90). The comparisons with tamoxifen and fulvestrant suggested no statistically significant differences.

Diarrhoea

Ten trials with 5200 women had data on diarrhoea toxicity. Of these, three compared an AI with tamoxifen, five with MA, and two with fulvestrant. AIs were associated with a statistically significant higher rate of diarrhoea than either tamoxifen (OR 1.64, 95% CI 1.06 to 2.55) or MA (OR 1.48, 95% CI 1.02 to 2.13) but not fulvestrant (P = 0.36).

Rash

Fifteen trials with 4598 women had data on rash toxicity. Of these, four compared an AI with tamoxifen, eight with MA, two with MPA, and one with fulvestrant. AIs were associated with a statistically significant increased risk of rash when compared with tamoxifen (OR 33.61, 95% CI 4.71 to 239.97) and for the two trials versus MPA (OR 36.80, 95% CI 3.35 to 404.73) but not against MA or fulvestrant. Within the comparison with MA there was statistically significant heterogeneity (P = 0.0003) and moderate heterogeneity with tamoxifen.

Vaginal bleeding Data on vaginal bleeding were reported in six trials of 2750 women: one compared an AI with tamoxifen, three with MA, and two with MPA. Compared with MA, there was a statistically significant benefit of 78% for treatment with the AI (OR 0.22, 95% CI 0.10 to 0.45). The two trials versus MPA also found a statistically significant difference with an OR of 0.13 (95% CI 0.02 to 0.71). In one of the larger trials (Bonneterre 2001) that compared an AI with tamoxifen, there was no statistically significant difference (P = 0.15).

Thromboembolic events

Thromboembolic event data were available from six trials with 2937 women. Two compared an AI with tamoxifen, three with MA, and one with fulvestrant. The AI had a statistically significant advantage only over tamoxifen (OR 0.48, 95% CI 0.27 to 0.85).

Arthralgia Data on arthralgia were available for 2470 women in two trials versus tamoxifen (1031 women) and four trials versus MA (1439 women). There was no statistically significant difference between the AIs and either tamoxifen or MA.

Subgroup analysis: aromatase inhibitors (AIs) in current clinical use versus any non-aromatase inhibitor

Eleven of the 32 trials comparing an AI with a non-AI were on the three AIs in current clinical use, namely anastrozole, exemestane and letrozole. The pooled results for these are reported. Data on overall survival and time to progression were available from only six and seven trials respectively, but response rates and clinical benefit were available from all 11 trials. In terms of survival, HRs were reported in the publications of four trials: anastrozole (Bonneterre 2001; Buzdar 1996a) and letrozole (Buzdar 2001; Dombernowsky 1998). Another two trials (Kaufmann 2000; Milla-Santos 2003) had sufficient data for calculation of the HRs. For time to progression, the corresponding numbers of trials were four (Bonneterre 2001; Buzdar 2001; Chia 2008; Mourisden 2001) and three (Dombernowsky 1998; Kaufmann 2000; Mauriac 2003), respectively.

1. Overall survival
Data on survival were available from six trials (Bonneterre 2001; Buzdar 1996a; Buzdar 2001; Dombernowsky 1998; Kaufmann 2000; Milla-Santos 2003). The AI was statistically significantly superior to the non-AI with a HR of 0.88 (95% CI 0.80 to 0.96), equivalent to a 12% benefit of treatment with an AI. This effect was consistent across all subgroups.

2. Progression-free survival
Data on progression were available from seven trials (Bonneterre 2001; Buzdar 2001; Chia 2008; Dombernowsky 1998; Kaufmann 2000; Mauriac 2003; Mourisden 2001) reporting an estimated 3660 events in 5004 women. Use of an AI was not statistically significantly associated with a change in the hazard of progression (HR 0.93, 95% CI 0.78 to 1.12). The results did not vary by type of AI. There was significant heterogeneity in the pooled result (P < 0.00001) within the anastrozole trials (P < 0.00001) and the letrozole trials (P = 0.01).

3. Proportion of women with clinical benefit (5619 assessable women)
Data were available from 11 trials. The pooled OR suggested a statistically significant advantage of 20% for the AI (OR 0.80, 95% CI 0.66 to 0.97). There was statistically significant heterogeneity among the trials (P = 0.002).

4. Proportion of women with an objective response (5619 assessable women)
All 11 trials reported objective response. The pooled OR of 0.79 (95% CI 0.65 to 0.97) showed a statistically significant advantage to the AI but there was statistically significant heterogeneity (P = 0.03) across the trial results. There was also significant heterogeneity within the exemestane trials.

5. Toxicity One of the suggested benefits of the third generation AIs is a reduced toxicity profile. The results were presented by comparator as the comparators have different toxicity profiles whereas the AIs have similar toxicity profiles. The denominators for the comparison of anastrozole with fulvestrant varied depending on whether the combined trial results were available (hot flushes, nausea, vomiting, thromboembolic events) or not (diarrhoea, rash).
Hot flushes Hot flushes were the specific toxicity that was reported most widely. Data on hot flushes were available from nine of the 11 trials, with 5623 women. Three trials compared the AI with tamoxifen, four with MA, and two with fulvestrant. The use of an AI had a very similar risk of hot flushes to tamoxifen and fulvestrant but was associated with statistically significant more reports of hot flushes than with MA (OR 1.69, 95% CI 1.24 to 2.30).
Nausea Data on nausea were available from nine of the 11 trials, involving 5623 women. Of the nine trials, three compared an AI with tamoxifen, four with MA and two with fulvestrant. The AIs had statistically signicantly more reports of nausea than MA (OR 1.45, 95% CI 1.09 to 1.95) but there was no statistically significant difference when the AIs were compared to tamoxifen or fulvestrant.

Vomiting
Five trials with 3499 women had data on vomiting alone and only MA as the comparator had more than one trial. There was no statistically significant differences between the AI and either tamoxifen or fulvestrant. Compared with MA, the AIs had a statistically significantly increased risk of vomiting (OR 1.77, 95% CI 1.11 to 2.83).

Diarrhoea Seven trials with 4295 women had data on diarrhoea toxicity. Two compared an AI with tamoxifen, three with MA, and two with fulvestrant. There was a statistically significant increased risk of diarrhoea with the AIs against MA (OR 2.40, 95% CI 1.34 to 4.29).

Rash Four trials with 2033 women that compared AIs with MA or fulvestrant (one trial only) had data on rash. AIs were not associated with a statistically significant increased risk of rash and there was statistically significant heterogeneity among the three trials with MA as the comparator (P = 0.04).
Vaginal bleeding Data on vaginal bleeding were reported in three trials with 1932 women, one compared an AI with tamoxifen and two with MA. There was a statistically significant benefit to treatment with the AIs in comparison with MA (OR 0.29, 95% CI 0.13 to 0.65).

Thromboembolic events Thromboembolic event data were available for 2378 women in three trials but there was only one trial per comparator (tamoxifen, MA, or fulvestrant). AIs were associated with a statistically significant lower incidence of thromboembolic events than tamoxifen (OR 0.53, 95% CI 0.30 to 0.96) but not compared with MA or fulvestrant.

Arthralgia
Data on arthralgia as a specific side effect were only available for 1394 women in three trials, two versus tamoxifen and one versus MA. Against both comparators, the AI was not statistically significantly associated with a difference in the incidence of arthralgia.

Other analyses
Aromatase inhibitors (AIs) versus any different aromatase inhibitor

A total of 2346 women in five trials were randomised to one AI versus a different AI. Of these, all five had data on response but only two had results on overall survival and progression-free survival (Gershanovich 1998; Goss 2007). Letrozole was compared with a different AI in all the trials (Gershanovich 1998; Rose 2003; Tominaga 2003) except that of Kleeberg 1997 which compared anastrozole with formestane. The trial by Rose and colleagues (Rose 2003) compared letrozole to anastrozole and in this section has been included in both the letrozole and anastrozole groups.
1. Overall survival
The Gershanovich 1998 and Goss 2007 trials were the only ones that had data on overall survival and the results were driven by Goss 2007 as 70% data came from this trial. Letrozole had a reduced HR of 0.91 (95% CI 0.82 to 1.02) but this was not statistically significant and there was signficant trial heterogeneity (P = 0.006).

2. Progression-free survival Two trials had data on progression from 1416 women (Gershanovich 1998; Goss 2007) and, again, the results were driven by the Goss 2007 trial. In these trials, letrozole was associated with a slightly reduced hazard in terms of progression-free survival compared to aminoglutethimide, but this was not statistically significant and there was heterogeneity (P = 0.01) between the trials.

3. Proportion of women with clinical benefit (1747 assessable patients) Letrozole was associated with a statistically significant clinical benefit compared with a different AI (OR 0.77, 95% CI 0.62 to 0.95). There was no significant trial heterogeneity (P = 0.63).

4. Proportion of women with an objective response (1747 assessable patients)
The pooled overall result was not presented as Rose 2003 was included in both individual AI comparisons and so would be counted twice. Letrozole was statistically significantly different from any other AI (OR 0.62, 95% CI 0.50 to 0.78). Results of all letrozole trials were consistent (test for heterogeneity P = 0.32). Anastrozole appeared to be significantly inferior to a different AI (OR 1.59, 95% CI 1.07 to 2.37).

Aromatase inhibition as first-line therapy versus any non-AI therapy (tamoxifen)

Twelve trials that randomised 3746 women used AIs exclusively as first-line therapy for advanced (metastatic) disease and all comparisons were against tamoxifen. We did not include any trials that were mixed first- and second-line. Data from three trials with 1483 women (anastrozole, fadrozole, AG) were available for overall survival and four trials with 2390 women (one trial each on formestane, anastrozole, and letrozole) for progression-free survival. Eleven trials reported results for objective response and nine trials for clinical benefit.

1. Overall survival
There was no statistically significant difference in the effect of treatment with an AI compared to tamoxifen.

2. Progression-free survival
The first-line AI regimen was statistically significantly superior to tamoxifen with a decreased hazard of 0.78 (95% CI 0.71 to 0.86). Anastrozole (Bonneterre 2001) and letrozole (Mourisden 2001) were statistically significantly different from tamoxifen (reduced hazard of 18% and 30%, respectively).

3. Proportion of women with clinical benefit (3252 assessable women) As results for individual AIs, except for aminoglutethimide and anastrozole, were based on only a single trial the pooled result is emphasised. The AIs were better than tamoxifen as first-line therapy (OR 0.69, 95% CI 0.51 to 0.92) although there was significant heterogeneity across the AIs (P = 0.002).
4. Proportion of women with objective response (3503 assessable women) Aminoglutethimide was the only AI with more than two trials published. The AIs were better than tamoxifen as first-line therapy (OR 0.77, 95% CI 0.59 to 1.00) although this was of borderline statistical signficance. There was considerable heterogeneity (P = 0.003) by type of AI. Exemestane and letrozole were the only AIs that were statistically significantly better than tamoxifen but in both cases the results were only based on one trial.
Aromatase inhibition as second-line therapy versus any non-AI therapy Women who had previously been treated with endocrine therapy, either a different AI or non-AI, for advanced (metastatic) disease and received the trial AI as second-line therapy were included in 19 trials. The trial by Leitzel 1995 was of second-line therapy but does not contribute to the results here, thus giving 18 trials. Aminoglutethimide was used as second-line in five trials, formestane in two, anastrozole in two, exemestane in two, fadrozole in three, letrozole in three, and vorozole in one trial. The majority of the comparisons (12) were against MA. We did not include trials where there was a mixture of first- and second-line therapy.

Data on objective response were available from all of the trials; clinical benefit from 16 trials; HRs for progression-free survival from eight trials; and HRs for overall survival from two trials.

1. Overall survival
Data on overall survival were limited with data from two trials of different AIs, anastrozole and letrozole. Second-line treatment with an AI was statistically significantly associated with a decreased hazard of death (HR 0.80, 95% CI 0.66 to 0.96). This effect was consistent for both AIs (heterogeneity P = 0.79).

2. Progression-free survival
AI use was not associated with a statistically significant difference in the risk of progression. There was significant heterogeneity (P = 0.001) across trials, with use of either anastrozole or vorozole associated with a significantly increased risk of progression.

3. Proportion of women with clinical benefit (5410 assessable women) There did not appear to be any effect in terms of a statistically significant clinical benefit when an AI was used as second-line therapy (OR 0.99, 95% CI 0.88 to 1.11). This lack of effect was consistent across AI subgroups (heterogeneity P = 0.88).

4. Proportion of women with objective response (5937 assessable women)
Overall there was no statistically significant difference between the use of an AI as second-line therapy and any other therapy (OR 0.98, 95% CI 0.86 to 1.13). When looking at individual AIs, none showed any evidence of a benefit but this was based on small numbers. There was no statistical heterogeneity (P = 0.52).

Other subgroup analysis

We were not able to perform subgroup analyses on the following groups of patients as these data were not systematically reported:

- ER positive versus ER unknown;
- according to site of distant metastases and differential treatment effect.

Quality of Life

Nine trials (Bezwoda 1998; Buzdar 1996b; Buzdar 1996c; Buzdar 2001; Chia 2008; Goss 1999; Kaufmann 2000; Mauriac 2003; Thuerlimann 1997) quoted quality of life (QOL) as a secondary endpoint. Three of the trials (Bezwoda 1998; Buzdar 1996b; Buzdar 1996c) did not report any QOL data. Only one (Thuerlimann 1997) has published two papers on the QOL data in detail. One trial (Dombernowsky 1998) mentioned that a QOL instrument was used, at baseline and at each visit whilst on treatment, but it was not mentioned as an endpoint nor were any data included. Chia 2008 reported that the difference in QOL between the treatment arms was not statistically significant; however the graph was shown on the online publication only.

There are several reasons why the limited QOL data are not included in this review: heterogeneous changes among patients, that is different symptoms and side effect profiles; different methods of drug application, that is injection versus tablets; use of four different QOL instruments at several different timepoints; some results given as responders versus non-responders rather than by treatment groups; some QOL measures based on clinician-reported rather than patient-reported symptoms.

DISCUSSION

Summary of main results

This review demonstrates a survival benefit of 10% with the use of AIs for the treatment of advanced (metastatic) breast cancer. This finding is not consistent across all AIs, with the greatest benefit (a survival benefit of 12%) associated with the AIs in current clinical use, namely anastrozole, exemestane and letrozole. However, data on survival were only available for about half the women and one of the trials (Buzdar 1996a) was not designed or powered to detect significant differences in survival.

The positive effects of AIs in terms of tumour response were statistically significant for first-line therapy where the comparator was tamoxifen. There were no data available for other comparators. When comparing the effect of the AI as second-line therapy there was no statistically significant difference on tumour response. In terms of progression-free survival, there was a statistically significant decrease in hazard of progression for treatment with the AIs as first-line therapy only. The paucity of data makes it difficult to make any firm conclusions in terms of overall survival.

In terms of toxicity, AIs are known to be associated with a higher incidence of nausea, diarrhoea, rash and arthralgia but a lower risk of vaginal bleeding and thromboembolic events. However, combining data across trials was difficult as both the toxicities reported and the criteria for reporting toxicities, if reported at all, varied greatly. We therefore did not have data on all predetermined toxicities with all comparators. Despite the inadequacies of the data, our review corroborated the direction of the known side effects. There was a higher incidence of hot flushes when compared to MA but not to tamoxifen; nausea compared to MA but not tamoxifen or fulvestrant; vomiting compared to MA but not tamoxifen or fulvestrant; diarrhoea compared to tamoxifen and MA but not fulvestrant; and rash compared to tamoxifen and MPA but not MA or fulvestrant. The risk of vaginal bleeding was about 80% lower with AI treatment and the incidence of thromboembolic events halved. For arthralgia, there was no statistically significant difference between the AIs and either tamoxifen or MA.

Overall completeness and applicability of evidence

A lack of standardised reporting of clinical endpoints impacted upon the analysis of all AIs, not just aminoglutethimide. Therefore, it was not possible to include all trials in each section. This reduced the power of certain analyses, especially overall and progression-free survival. In addition, many of the data required to carry out analyses of prospectively identified subgroups, as set out in the review protocol, were not available. We could not, therefore, identify specific subgroups of women who may benefit from AI use.

There are very limited data on quality of life reported in this setting. The limited quality of life data which was reported did not show any significant differences between the AI and comparator groups; however, some differences were found with some subscales in favour of the AI (Goss 1999; Kaufmann 2000). The patient's perspective in advanced disease treatment is an important endpoint and should be included in trials as it would aid interpretation in this mainly palliative setting.

Quality of the evidence

This review has combined data from a wide variety of trials that were carried out over 20 years. Some of the trials did not use an AI as a single agent but in combination with another endocrine therapy. There was heterogeneity both across types of AI and within each AI. The results of trials of three generations of AIs have been combined as well as results from trials of steroidal and non-steroidal therapy. This has been forced, to some extent, by the lack of data on individual AIs.

Within each AI, trials varied in terms of sample size, dose of AI, comparison regimen, outcomes, length of follow up and quality of reporting. For example, the 11 trials of aminoglutethimide consisted of between 62 and 313 patients; four of the trials were of first-line therapy, five second-line, and two mixed. Doses of aminoglutethimide used were 125 mg in one trial, 250 mg* in three, 500 mg* in four, 750 mg in one, and 1000 mg in two (* dose doubled after a specific period of treatment). The comparator was tamoxifen in five trials (20 mg in three, 30 mg in one, 40 mg in one), MA 160 mg in three trials, MPA 500 mg in one trial, MPA 1000 mg in four trials, and HC 20 mg in one trial. Not all endpoints were available in each trial and four reported overall survival, three progression-free survival, eight clinical benefit, and 10 objective response.

Potential biases in the review process

If the description of randomisation is used as a barometer of reporting trial quality, it appears that this has improved over time. For example, in the trials of the first generation AI aminoglutethimide six of 11 randomisations were categorised as unclear whereas only two of the nine trials of third generation AI letrozole were considered as such.

Evidence of heterogeneity between trials was identified for tumour response rates and progression-free survival though not overall survival. The reasons for this are unknown but this statistical heterogeneity might be explained by clinical heterogeneity. It may be that outcomes involving the subjective endpoint, that is tumour response, are subject to variation whereas the hard endpoint used in the survival analysis is unequivocal. Other contributory factors may be the difference in dosage of some AIs and significant differences in the proportion of patients who were truly hormone receptor positive.

Agreements and disagreements with other studies or reviews

In September 2006, Mauri and colleagues published a paper entitled, "Survival with aromatase inhibitors and inactivators versus standard hormonal therapy in advanced breast cancer: meta-analysis" Mauri 2006 which came to the same conclusion; that is, "Inhibition of the AI system, in particular with third generation AIs, appears to be associated with statistically significant improved survival of patients with advance breast cancer compared with standard hormonal treatments".

AUTHORS' CONCLUSIONS

Implications for practice

Historically, the treatment for advanced (metastatic) breast cancer has been with hormonal treatments such as tamoxifen or the progestins MA or MPA. This review confirms a survival benefit

of treating advanced (metastatic) breast cancer with the third generation aromatase inhibitors (anastrozole, exemestane, and letrozole) that are being used clinically today.

Implications for research

This review would benefit from additional publications with greater survival details, that is median survival and number of events, for those trials that did not publish them originally. Further data from exemestane trials are required to evaluate this AI more completely. Efforts should be made to standardise reporting of toxicity and a quality of life component should also be included.

ACKNOWLEDGEMENTS

The authors were funded by Cancer Research UK. Lorna Gibson would like to thank Julian Peto and Isabel dos Santos Silva for their support. The authors would like to thank the Cochrane Breast Cancer Group for their support and are indebted to Olivia Fletcher for critical reading of the manuscript.

REFERENCES

References to studies included in this review

Alonso-Munoz 1988 *{published data only}*

* Alonso-Munzo MC, Ojeda-Gonzalez M, Beltran-Fabregat M, Dorca-Ribugent J, Lopez-Lopez L, Borras-Balada J et al. Randomized trial of tamoxifen versus aminoglutethimide and versus combined tamoxifen and aminoglutethimide in advanced postmenopausal breast cancer. *Oncology* 1988;**45**:350-3.

Bezwoda 1998 *{published data only}*

* Bezwoda WR, Gudgeon A, Falkson G, Jordaan JP, Goedhals L. Fadrozole versus megestrol acetate: a double-blind randomised trial in advanced breast cancer. *Oncology* 1998;**55**:416-20.

Bonneterre 2001 *{published data only}*

* Bonneterre J, Buzdar A, Nabholtz JA, Robertson JFR, Thuerlimann B, von Euler M et al. Anastrozole is superior to tamoxifen as first-line therapy in hormone receptor positive advanced breast carcinoma. *Cancer* 2001;**92**(9):2247-58.

Bonneterre J, Thuerlimann B, Robertson JFR, Krzakowski M, Mauriac L, Koralewski P, et al. Anastrozole versus tamoxifen as first-line therapy for advanced breast cancer in 668 postmenopausal women: results of the tamoxifen or arimidex randomized group efficacy and tolerability study. *Journal of Clinical Oncology* 2000;**18**(22):3748-57.

Nabholtz JM, Bonneterre J, Buzdar A, Robertson JFR, Thuerlimann B. Anastrozole (Arimidex) versus tamoxifen as first-line therapy for advanced breast cancer in postmenopausal women: survival analysis and updated safety results. *European Journal of Cancer* 2003;**39**:1684-9.

Nabholtz JM, Buzdar A, Pollak M, Harwin W, Burton G, Mangalik A et al. Anastrozole is superior to tamoxifen as first-line therapy for advanced breast cancer in postmenopausal women: results of a North American multicenter randomized trial. *Journal of Clinical Oncology* 2000;**18**(22):3758-67.

Nabholtz JM. Advanced breast cancer updates on anastrozole versus tamoxifen. *Journal of Steroid Biochemistry and Molecular Biology* 2003;**86**:321-5.

Thuerlimann B, Robertson JFR, Nabholtz JM, Buzdar A, Bonneterre J. Efficacy of tamoxifen following anastrozole ('Arimidex') compared with anastrozole following tamoxifen as first-line treatment for advanced breast cancer in postmenopausal women. *European Journal of Cancer* 2003;**39**:2310-17.

Vergote I, Bonneterre J, Thuerlimann B, Robertson J, MKrzakowski M, Mauriac L et al. Randomised study of anastrozole versus tamoxifen as first-line therapy for advanced breast cancer in postmenopausal women. *European Journal of Cancer* 2000;**36**(Suppl:):81-91.

Buzdar 1996a *{published data only}*

* Budzar A, Jonat W, Howell A, Jones SE, Blomqvist C, Vogel CL et al. Anastrozole, a potent and selective aromatase inhibitor, versus megestrol acetate in postmenopausal women with advanced breast cancer: results of overview analysis of two phase III trials. *Journal of Clinical Oncology* 1996;**14**:2000-11.

Buzdar AU, Jonat W, Howell A, Jones SE, Blomqvist CP, Vogel CL et al. Anastrozole versus megestrol acetate in the treatment of postmenopausal women with advanced breast carcinoma. *Cancer* 1998;**83**(6):1142-52.

Buzdar AU, Jonat W, Howell A, Plourde PV. Arimidex: a potent and selective aromatase inhibitor for the treatment of advanced breast cancer. *Journal Steroid Biochemistry Molecular Biology* 1997;**61**(3-6):145-9.

Drummond M, Thompton E, Howell A, Jonat W, Buzdar A, Brown J. Cost-effectiveness implications of increased survival with anastrozole in the treatment of advanced breast cancer. *Journal of Medical Economics* 1999;**2**:33-43.

Jonat W, Howell A, Blomqvist C, Eiermann W, Winblad G, TyrrellC et al. A randomized trial comparing two doses of the new selective aromatase inhibitor anastrozole (Arimidex) with megestrol acetate in postmenopausal women with advanced breast cancer. *European Journal of Cancer* 1996;**32A**(3):404-12.

Robertson JFR, Howell A, Buzdar A, von Euler M, Lee D. Static disease on anastrozole provides similar benefit as objective response in patients with advanced breast cancer. *Breast Cancer Research and Treatment* 1999;**58**:157-62.

Buzdar 1996b *{published data only}*

* Buzdar A, Smith R, Vogel C, Bonomi P, Keller AM, Favis G et al. Fadrazole HCl (CGS-16949A) versus megestrol acetate in postmenopausal patients with metastatic breast carcinoma: results of two randomized, double-blind, controlled multi-institutional trials. *Cancer* 1996;**77**(12):2503-13.

Buzdar 1996c *{published data only}*

Buzdar A, Smith R, Vogel C, Bonomi P, Keller AM, Favis G et al. Fadrazole HCl (CGS-16949A) versus megestrol acetate in postmenopausal patients with metastatic breast carcinoma: results of two randomized, double-blind, controlled multi-institutional trials. *Cancer* 1996;**77**(12):2503-13.

Buzdar 2001 *{published data only}*

* Buzdar A, Douma J, Davidson N, Elledge R, Morgan M, Smith R et al. Phase III, multicenter, double-blind, randomized study of letrozole, an aromatase inhibitor, for advanced breast cancer versus megestrol acetate. *Journal of Clinical Oncology* 2001;**19**(14):3357-66.

Canney 1988 *{published data only}*

* Canney PA, Priestman TJ, Griffiths T, Latief TN, Mould JJ, Spooner D. Randomized trial comparing aminoglutethimide with high-dose medroxyprogesterone acetate in therapy for advanced breast carcinoma. *Journal of the National Cancer Institute* 1988;**80**(14):1147-51.

Chia 2008 *{published data only}*

* Chia S, Gradishar W, Mauriac L, Bines J, Amant F, Federico M et al. Double-blind, randomized placebo controlled trial of

fulvestrant compared with exemestane after prior nonsteroidal aromatase inhibitor therapy in postmenopausal women with hormone receptor-positive, advanced breast cancer: results from EFECT. *Journal of Clinical Oncology* 2008;**26**(10):1664-70.

Dombernowsky 1998 *{published data only}*

* Dombernowsky P, Smith I, Falkson G, Leonard R, Panasci L, Bellmunt J et al. Letrozole, a new oral aromatase inhibitor for advanced breast cancer: double-blind randomized trial showing a dose effect and improved efficacy and tolerability compared with megestrol acetate. *Journal of Clinical Oncology* 1998;**16**(2):453-61.

Falkson 1996 *{published data only}*

* Falkson CI, Falkson HC. A randomised study of CGS 16949A (fadrozole) versus tamoxifen in previously untreated postmenopausal patients with metastatic breast cancer. *Annals of Oncology* 1996;**7**:465-9.

Freue 2000 *{published data only}*

Freue M, Kjaer C, Boni J, Joliver F, Jaenicke F, Willemse PhHB et al. Open comparative trial of formestane versus megestrol acetate in postmenopausal patients with advanced breast cancer previously treated with tamoxifen. *The Breast* 2000;**9**:9-16.

Gale 1994 *{published data only}*

Gale KE, Andersen JW, Tormey DC, Mansour EG, David TE, Horton J et al. Hormonal treatment for metastatic breast cancer. An Eastern Cooperative Oncology Group Phase III trial comparing aminoglutethimide to tamoxifen. *Cancer* 1994;**73**(2):354-361.

Garcia-Giralt 1992 *{published data only}*

Garcia-Giralt E, Ayme Y, Carton M, Daban A, Delozier T, Fargeot P et al. Second and third line hormonotherapy in advanced post-menopausal breast cancer: a multicenter randomized trial comparing medroxyprogesterone acetate with aminoglutethimide in patients who have become resistant to tamoxifen. *Breast Cancer Research and Treatment* 1992;**24**:139-45.

Gershanovich 1998 *{published data only}*

* Gershanovich M, Chaudri HA, Campos D, Lurie H, Bonaventura A, Jeffrey M et al. Letrozole, a new oral aromatase inhibitor: randomised trial comparing 2.5mg daily, 0.5 mg daily and aminoglutethimide in postmenopausal women with advanced breast cancer. *Annals of Oncology* 1998;**9**:639-45.

Gershanovitch M. Letrozole, a new aromatase for advanced breast cancer. *Voprosy Onkolgii* 45;**4**:361-8.

Goss 1999 *{published data only}*

* Goss PE, Winer EP, Tannock IF, Schwarz LH. Randomized phase III trial comparing the new potent and selective third generation aromatase inhibitor vorozole with megestrol acetate in postmenopausal advanced breast cancer patients. *Journal of Clinical Oncology* 1999;**17**(1):52-63.

Goss 2007 *{published data only}*

Goss P, Bondarenko IN, Mankhas GN, Pendergrass KB, Miller WH, Langecker P et al. Phase III, double-blind, controlled trial of atamastane plus toremifene compared with letrozole in postmenopausal women with advanced receptor-positive breast cancer. *Journal of Clinical Oncology* 2007;**25**(31):4961-6.

Ingle 1986 *{published data only}*

* Ingle JN, Green SJ, Ahmann DL, Long HJ, Edmonson JH, Rubbin J et al. Randomized trial of tamoxifen alone or combined with aminoglutethimide and hydrocortisone in women with metastatic breast cancer. *Journal of Clinical Oncology* 1986;**4**(6):958-64.

Kaufmann 2000 *{published data only}*

Hillner BE, Radice D. Cost-effectiveness analysis of exemestane compared with megestrol in patients with advanced breast carcinoma. *Cancer* 2001;**91**(3):484-9.

Kaufmann M, Bajetta E, Dirix LY, Fein LE, Jones SE, Cervek J et al. Exemestane improves survival compared with megestrol acetate in postmenopausal patients with advanced breast cancer who have failed on tamoxifen: results of a double-blind randomised phase III trial. *European Journal of Cancer* 2000;**36 Suppl**:81-91.

* Kaufmann M, Bajetta E, Dirix LY, Fein LE, Jones SE, Zilembo N et al. Exemestane is superior to megestrol acetate after tamoxifen failure in postmenopausal women with advanced breast cancer: results of a phase III randomised double-blind trial. *Journal of Clinical Oncology* 2000;**18**(7):1399-411.

Kleeberg 1997 *{published data only}*

* Kleeberg UR, Dowsett M, Carrion RP, Dodwell DJ, Vorobiof DA, Aparicio LA, Robertson JFR. A randomised comparison of oestrogen suppression with anastrozole and formestane in postmenopausal patients with advanced breast cancer. *Oncology* 1997;**54 Suppl 2**:19-22.

Vorobiof DA, Kleeberg UR, Perez-Carrion R, Dodwell DJ, Robertson JFR, Calvo L et al. A randomized, open, parallel-group trial to compare the endocrine effects of oral anastrozole (Arimidex) with intramuscular formestane in postmenopausal women with advanced breast cancer. *Annals of Oncology* 1999;**10**:1219-25.

Leitzel 1995 *{published data only}*

* Leitzel K, Teramoto Y, Konrad K, Chinchilli VM, Volas G, Grossberg H et al. Elevated Serum c-erb B-2 Antigen levels and decreased response to hormone therapy of breast cancer. *Journal of Clinical Oncology* 1995;**13**(5):1129-35.

Lundgren 1989 *{published data only}*

Lundgren S, Gundersen S, Klepp R, Lonning PE, Lund E, Kvinnsland S. Megestrol acetate versus aminoglutethimide for metastatic breast cancer. *Breast Cancer Research and Treatment* 1989;**14**:201-6.

Mauriac 2003 *{published data only}*

Howell A, Robertson JFR, Quaresma Albano J, Ascgermannova A, Mauriac L, Kleeberg UR et al. Fulvestrant, formerly ICI 182, 780, is as effective as anastrozole in

postmenopausal women with advanced breast cancer progressing after prior endocrine treatment. *Journal of Clinical Oncology* 2002;**20**(16):3396-403.

* Mauriac L, Pippen JE, Quaresma Albano J, Gertler SZ Osborne CK. Fulvestrant (Faslodex) versus anastrozole for the second-line treatment of advanced breast cancer in subgroups of postmenopausal women with visceral and non-visceral metasteses: combined results from two multicentre trials. *European Journal of Cancer* 2003;**39**(9):1228-33.

Osborne CK, Pippen J, Jones, SE, Parker LM, Ellis M, Come S et al. Double-blind, randomized trial comparing the efficacy and tolerability of the fulvestrant versus anastrozole in postmenopausal women with advanced breast cancer progressing on prior endocrine therapy: results of a North American Trial. *Journal of Clinical Oncology* 2002;**20**(16):3386-95.

Robertson JFR, Osborne CK, Howell A, Jones SE, Mauriac L, Ellis M et al. Fulvestrant versus anastrozole for the treatment of advanced breast carcinoma in postmenopausal women. *Cancer* 2003;**98**(2):229-38.

Vergote I. Fulvestrant versus anastrozole as second-line treatment of advanced breast cancer in postmenopausal women. European Journal of Cancer 2002;**38 Suppl**(6):57-8.

Mercer 1993 *{published data only}*

* Mercer PM, Ebbs SR, Fraser SCA, Coltart RS, Bates T. Trial of aminoglutethimide vs hydrocortisone as second-line hormone treatment of advanced breast cancer. *European Journal of Surgical Oncology* 1993;**19**:254-8.

Milla-Santos 2003 *{published data only}*

Milla-Santos A, Milla L, Portella J, Rallo L, Pons M, Rodes E et al. Anastrozole versus tamoxifen as first-line therapy in postmenopausal patients with hormone-dependent advanced breast cancer. *American Journal of Clinical Oncology* 2003;**26**(3):317-22.

Mourisden 2001 *{published data only}*

Irish W, Sherrill B, Cole B, Gard C, Glendenning GA, Mourisden H. Quality-adjusted survival in a crossover trial of letrozole versus tamoxifen in postmenopausal women with advance breast cancer. *Annals of Oncology* 2005;**16**:1458-62.

Lipton A, Ali SM, Leitzel K, Demers L, Harvey HA, Chaudri-Ross HA et al. Serum HER-2/neu and response to the aromatase inhibitor letrozole versus tamoxifen. *Journal of Clinical Oncology* 2003;**21**(10):1967-72.

Mourisden H, Chaudri-Ross HA. Efficacy of First-Line letrozole versus tamoxifen as a function of age in postmenopausal women with advanced breast cancer. *The Oncologist* 2004;**9**:497-506.

Mourisden H, Gershanovich M, Sun Y, Perez-Carrion R, Boni C, Monnier A et al. Phase III study of letrozole versus tamoxifen as first-line therapy of advanced breast cancer in postmenopausal women: analysis of survival and update of efficacy from the International Letrozole Breast Cancer Group. *Journal of Clinical Oncology* 2003;**21**(11):2101-9.

* Mourisden H, Gershanovich M, Sun Y, Perez-Carrion R, Boni C, Monnier A et al. Superior efficacy of letrozole versus tamoxifen as first-line therapy for postmenopausal women with advanced brreast cancer: results of a phase III study of the international letrozole breast cancer group. *Journal of Clinical Oncology* 2001;**19**(10):2596-606.

Mourisden H, Sun Y, Gershanovich M, Perez-Carrion R, Becquart D, Chaudri-Ross HA, Lang R. Superiority of letrozole to tamoxifen in the first-line treatment of advanced breast cancer: evidence from metastatic subgroups and a test of functional ability. *The Oncologist* 2004;**9**:489-96.

Mourisden HT. Letrozole in advanced breast cancer: the PO25 trial. *Breast Cancer Research and Treatment* 2007;**105**:19-29.

Paridaens 2003 *{published data only}*

Atalay G, Dirix L, Biganzoli L, Beex L, Nooji M, Cameron D et al. The effect of exemestane on serum lipid profile in postmenopausal women with metastatic breast cancer: a comparison study to EORTC Trial 10951, 'Randomized phase II study in first line hormonal treatment for metastatic breast cancer with exemestane or tamoxifen in postmenopausal patients'. *Annals of Oncology* 2004;**15**(2):211-7.

* Paridaens R, Dirix L, Lohrisch C, Beex L, Nooji M, Cameron D et al. Mature results of a randomized phase II multicenter study of exemextane versus tamoxifen as first-line hormone therapy for postmenopausal women with metastatic breast cancer. *Annals of Oncology* 2003;**14**:1391-8.

Perez Carrion 1994 *{published data only}*

* Perez Carrion R, Alberola Candel V, Calabresi F, Michel RT, Santos R, Delozier T et al. Comparison of the selective aromatase inhibitor formestane with tamoxifen as first-line therapy in postmenopausal women with advanced breast cancer. *Annals of Oncology* 1994;**5 Suppl 7**:19-24.

Powles 1984 *{published data only}*

* Powles TJ, Ashley S, Ford HT, Gazet JC, Nash AG, Neville AM, Coombes RC. Treatment of disseminated breast cancer with tamoxifen, aminoglutethimide, hydrocortisone, and danazol, used in combination or sequentially. *The Lancet* 1984;**1**(8391):1369-72.

Rose 1986 *{published data only}*

Rose C, Kamby C, Mouridsen HT, Andersson M, Bastholt L, Moller KA et al. Combined endocrine treatment of elderly postmenopausal patients with metastatic breast cancer. *Breast Cancer Research and Treatment* 2000;**61**(2):103-10.

* Rose C, Kamby C, Mouridsen HT, Bastholt L, Brincker H, Skovaard-Poulsen H et al. Combined endocrine treatment of postmenopausal patients with advanced breast cancer. A randomised trial of tamoxifen vs. tamoxifen plus aminoglutethimide and hydrocortisone. *Breast Cancer Research and Treatment* 1986;**7 Suppl**:45-50.

Rose 2003 *{published data only}*

* Rose C, Vtoraya O, Pluzanska A, Davidson N, Gershanovich M, Thomas R et al. An open randomised trial of second-line

endocrine therapy in advanced breast cancer. *European Journal of Cancer* 2003;**39**(16):2318-27.

Tobias JS, Howell A. An open randomised trial of second-line endocrine therapy in advanced breast cancer: comparison of the aromatase inhibitors letrozole and anastrozole. *European Journal of Cancer* 2004;**40**(12):1913.

Russell 1997 *{published data only}*

* Russell C, Green SJ, O'Sullivan J, Hynes HE, Budd T, Congdon JE et al. Megestrol acetate and aminoglutethimide/hydrocortisone in sequence or in combination as second-line endocrine therapy of estrogen receptor-positive metastatic breast cancer: a southwest oncology group phase III trial. *Journal of Clinical Oncology* 1997;**15**(7):2494-501.

Samonis 1994 *{published data only}*

Samonis G, Margioris AN, Bafaloukos D, Razis DV. Prospective randomized study of aminoglutethimide (AG) versus medroxyprogesterone acetate (MPA) versus AG+MPA in generalized breast cancer. *Oncology* 1994;**51**(5):411-5.

Schmid 2001 *{published data only}*

* Schmid P, Wischnewsky MB, Possinger K. Self organizing maps and prognosis of advanced breast cancer patients with bone metastases receiving letrozole or MA. *Breast Cancer Research and Treament* 2001;**64**:77.

Thuerlimann 1996 *{published data only}*

* Thuerlimann B, Beretta K, Bacchi M, Castiglione-Gertsch M, Goldhirsch A, Jungi WF et al. First-line fadrozole HCl (CGS 16949A) versus tamoxifen in postmenopausal women with advanced breast cancer. *Annals of Oncology* 1996;**7**:471-9.

Thuerlimann 1997 *{published data only}*

Bernhard J, Casiglione-Gertsch M, Hsu Schmitz S-F, Thuerlimann B, Cavalli F, Morant R et al. Quality of life in postmenopausal patients with breast cancer after failure of tamoxifen: formestane versus megestrol acetate as second-line hormonal treatment. *European Journal of Cancer* 1999;**35**(6):913-20.

Bernhard J, Thuerlimann B, Hsu Schmitz S-F, Castiglione-Gertsch M, Cavalli F, Morant R et al. Defining clinical benefit in postmenopausal patients with breast cancer under second-line endocrine treatment: does quality of life matter. *Journal of Clinical Oncology* 1999;**17**(6):1672-9.

* Thuerlimann B, Castiglione M, Hsu-Scmitz SF, Cavalli F, Bonnefoi H, Fey MF et al. Formestane versus megestrol acetate in postmenopausal breast cancer patients after failure of tamoxifen: a phase III prospective randomised cross over trial of second-line hormonal treatment (SAKK 20/90). *European Journal of Cancer* 1997;**33**(7):1017-24.

Tominaga 2003 *{published data only}*

* Tominaga T, Adachi I, Sasaki Y, Tabei T, Ikeda T, Takatsuka Y et al. Double-blind randomised trial comparing the non-steroidal aromatase inhibitors letrozole and fadrozole in postmenopausal women with advanced breast cancer. *Annals of Oncology* 2003;**14**:62-70.

References to studies excluded from this review

Abe 2002 *{published data only}*

* Abe R, Tominaga T, Nomizu T, Nomura Y, Takashima S, Koyama H et al. CGS20267 (letrozole), a new aromatase inhibitor: late phase II study for postmenopausal women with advanced or recurrent breast cancer (no. 1) - investigation of recommended clinical dose CGS20267 Study Group. *Gan To Kaguku Ryoho* 2002;**29**(5):729-40.

Bajetta 1994 *{published data only}*

* Bajetta E, Zilembo N, Buzzoni R, Noberasco C, Di Leo A, Bartoli C et al. Endocrinological and clinical evaluation of two doses of formestane in advanced breast cancer. *British Journal of Cancer* 1994;**70**:145-50.

Ferrari L, Zilembo N, Bajetta, Buzzoni R, Noberasco C, Martinetti A et al. Effect of two 4-hydroxyandrostenedione doses on serum insulin-like growth factor I levels in advanced breast cancer. *Breast Cancer Research and Treatment* 1994;**30**:127-32.

Bajetta 1997 *{published data only}*

* Bajetta E, Zilembo N, Noberasco C, Martinetti A, Mariani L, Ferrari L et al. The minimal effective exemestane dose for endocrine activity in advanced breast cancer. *European Journal of Cancer* 1997;**33**(4):587-91.

Bajetta 1997a *{published data only}*

* Bajetta E, Ferrari L, Celio L, Mariana L, Miceli R, Di Leo A et al. The aromatase inhibitor letrozole in advanced breast cancer: effects on serum insulin-like growth factor (IGF)-I and IGF-binding protein-3 levels. *Journal of Steroid Biochemistry* 1997;**63**(4-6):261-7.

Bajetta 1999 *{published data only}*

* Bajetta E, Zilembo N, Dowsett M, Guillevin L, Di Leo A, Celio L et al. Double-blind, randomised, multicentre endocrine trial comparing two letrozole doses in postmenopausal breast cancer patients. *European Journal of Cancer* 1999;**35**(2):208-13.

Beretta 1990 *{published data only}*

* Beretta KR, Hoeffken K, Kvinnsland S, Trunet P, Chaudri HA, Bhatnagar AS et al. CGS 16949A, a new aromatase inhibitor in the treatment of breast cancer - a phase I study. *Annals of Oncology* 1990;**1**(6):421-6.

Bruning 1989 *{published data only}*

* Bruning PF, Bonfer JMG, Hart AAM, van der Linden E, de Jong-Bakker M, Moolenaar AJ, Nooijen WJ. Low dose aminoglutethimide without hydrocortisone for the treatment of advanced postmenopausal breast cancer. *European Journal of Clinical Oncology* 1989;**25**(2):369-76.

Bruning 1990 *{published data only}*

* Bruning PF, Bonfrer JMG, Wildiers J, Jassem J, Beex LVAM, Schorangel J, Nooijen WJ. Second-line endocrine treatment of postmenopausal advanced breast cancer preliminary endocrine results of a 5-arm randomized phase II trial of medium vs low dose aminoglutethimide, with or without hydrocortisone vs hydrocortisone alone (EORTC 10861). *Journal of Steroid Biochemistry* 1990;**37**(6):1013-9.

Castelazo 2004 *{published data only}*

Castelazo RG, Molotia Xolalpa D, Basavilvazo Rodriguez MA, Angeles Victoria L, Zarate A, Hernandez-Valencia M. Survival of breast cancer patients treated with inhibitors of the aromatase vs tamoxifen. *Ginecologia y Obstetricia de Mexico* 2004;**72**:493-9.

Cataliotti 2006 *{published data only}*

* Cataliotti L, Buzdar AU, Noguchi S, Bines J, Takatsuka Y, Petrakova K. Comparison of anastrozole versus tamoxifen as preoperative therapy in postmenopausal women with hormone receptor-positive breast cancer. *Cancer* 2006;**106**(10):2095-103.

Dixon 2000 *{published data only}*

* Dixon JM, Renshaw L, Bellamy C, Stuart M, Hoctin-Boes G, Miller WR. The effects of neoadjuvant anastrozole (Arimidex) on tumor volume in postmenopausal women with breast cancer: a randomized, double-blind, single-center study. *Clinical Cancer Research* 2000;**6**(6):2229-35.

Dowsett 1989 *{published data only}*

Dowsett M, Cunningham DC, Stein RC, Evans S, Dehennin L, Hedley A, Coombes RC. Dose-related endocrine effects and pharmokinetics of oral and intramuscular 4-hydroxyandrostenedione in postmenopausal breast cancer patients. *Cancer Research* 1989;**49**:1306-12.

Dowsett 1990 *{published data only}*

* Dowsett M, Stein RC, Mehta A, Coombes RC. Potency and selectivity of the non-steroidal aromatase inhibitor CGS 16949A in postmenopausal breast cancer patients. *Clinical Endocrinology* 1990;**32**:623-34.

Dowsett 1994 *{published data only}*

Bonnefoi HR, Smith IE, Dowsett M, Trunet, Houston SJ, da Luz RJ et al. Therapeutic effects of the aromatase inhibitor fadrozole hydrochloride in advanced breast cancer. *British Journal of Cancer* 1996;**73**:539-42.

* Dowsett M, Smithers D, Moore J, Trunet PF, Coombes RC, Powles TJ et al. Endocrine changes with the aromatase inhibitor fadrozole hydrochloride in breast cancer. *European Journal of Cancer* 1994;**30A**(10):1453-8.

Dowsett 1995 *{published data only}*

* Dowsett M, Jones A, Johnston SRD, Jacobs S, Trunet P, Smith IE. In vivo measurement of aromatase inhibition by letrozole (CGS 20267) in postmenopausal patients with breast cancer. *Clinical Cancer Research* 1995;**1**:1511-5.

Eiermann 2001 *{published data only}*

* Eiermann W, Paepke S, Appfelstaedt J, Llombart-Cussac A, Eremin J, Vinholes J et al. Preoperative treatment of postmenopausal breast cancer patients with letrozole: a randomized double-blind multicenter study. *Annals of Oncology* 2001;**12**:1527-32.

Geisler 1996 *{published data only}*

* Geisler J, King N, Dowsett M, Ottestad L, Lundgren S, Walton P et al. Influence of anastrozole (Arimidex), a selective, non-steroidal aromatase inhibitor, on in vivo aromatisation and plasma oestrogen levels in post menopausal women with breast cancer. *British Journal of Cancer* 1996;**74**:1286-91.

Geisler 2002 *{published data only}*

* Geisler J, Haynes B, Anker G, Dowsett M, Loenning PE. Influence of letrozole and anastrozole on totla body aromatization and plasma estrogen levels in postmenopausal breast cancer patients evaluated in a randomised, cross-over study. *Journal of Clinical Oncology* 2002;**20**(3):751-7.

Ingle 1997 *{published data only}*

* Ingle JN, Johnson PA, Suman VJ, Gerstner JB, Mailliard JA, Camoriano JK et al. A randomized phase II trial of two dosage levels of letrozole as third-line hormonal therapay for women with metastatic breast carcinoma. *Cancer* 1997;**80**:218-24.

Johnston 1994 *{published data only}*

* Johnston SRD, Smith IE, Doody D, Jacobs S, Robertshaw H, Dowsett M. Clinical and endocrine effects of the oral aromatase inhibitor vorozole in postmenopausal patients with advanced breast cancer. *Cancer Research* 1994;**54**:5875-81.

Miller 1996b *{published data only}*

* Miller AA, Lipton A, Henderson IC, Navari R, Mulagha MT, Cooper J. Fadrozole hydrochloride in postmenopausal patients with metastatic breast carcinoma. *Cancer* 1996;**78**:789-93.

Pronzato 1993 *{published data only}*

* Pronzato P, Rubagotti A, Amoroso D, Bertelli G, Queirolo P, Sertoli MR, Rosso M et al. Second-line hormone therapy for breast cancer. Uselessness of first-line continuation. *American Journal of Clinical Oncology* 1993;**16**(6):522-5.

Raats 1992 *{published data only}*

Falkson G, Raats JI, Falkson H. Fadrozole hydrochloride, a new nontoxic aromatase inhibitor for the treatment of patients with metastatic breast cancer. *Journal of Steroid and Biochemistry and Molecular Biology* 1992;**43**(1-3):161-5.

* Raats JI, Falkson G, Falkson H. A study of fadrozole, a new aromatase inhibitor, in postmenopasual women with advanced metastatic breast cancer. *Journal of Clinical Oncology* 1992;**10**(1):111-6.

Smith 2005 *{published data only}*

* Smith IE, Dowsett M, Ebbs SR, Dixon MJ, Skene A, Blohmer JU et al. Neoadjuvant treatment of postmenopausal breast cancer with anastrozole, tamoxifen, or both in combination: the IMPACT multicenter double-blind randomized trial. *Journal of Clinical Oncology* 2005;**23**(22):5108-16.

Svenstrup 1994 *{published data only}*

* Svenstrup B, Herrstedt J, Bruenner N, Bennet P, Wachmann H, Dombernowsky P. Sex hormone levels in postmenopausal women with advanced metastatic breast cancer treated with CGS 169 49A. *European Journal of Cancer* 1994;**30A**(9):1254-8.

Wang 2003 *{published data only}*

* Wang HQ, Ren CC, Feng WJ, Lin MB, Xiong YQ, Huang P et al. Treatment of advanced metastatic breast cancer with

exemestane, a multicenter randomized controlled study of 195 cases. *Zhonghua Yi Xue Za Zhi* 2003;**83**(3):188-90.

References to ongoing studies

CAAN *{published data only}*

CAAN. Ongoing study. February 2002. Contact author for more information.

ECOG E4101 *{published data only}*

ECOG E4101. Ongoing study. Starting date of trial not provided. Contact author for more information.

ICR-CTSU Sofea *{published data only}*

SofeaPhase III. Ongoing study. March 2004. Contact author for more information.

Paridaens 2003 *{published data only}*

Phase III EORTC-10951. Ongoing study. Starting date of trial not provided. Contact author for more information.

Additional references

Beatson 1896

Beatson GT. On the treatment of inoperable cases of carcinoma of the mamma. Suggestions for a new method of treatment with illustrative cases. *Lancet* 1896;**2**:104-7.

EBCTCG 2005

Early Breast Cancer Trialists' Cooperative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet* 2005;**365**:1687-1717.

Ferlay 2000 [Computer program]

Globocan 2000: Cancer Incidence, Mortality and Prevalence Worldwide. Ferlay J, Bray F, Pisani P and Parkin DM, Version 1. Lyon: IARC Press, 2001.

Gibson 2007

Gibson LJ, Dawson CL, Lawrence DJ, Bliss JM. Aromatase inhibitors for treatment of advanced breast cancer in postmenopausal women. *Cochrane Database of Systematic Reviews* 2007, Issue 1. [DOI: 10.1002/14651858]

Hayward 1977

Hayward JL, Carbone PP, Heuson JC, Kumaoka S, Segaloff A, Rubens RD. Assessment of response to therapy in advanced breast cancer: a project of the Programme on Clinical Oncology of the International Union Against Cancer, Geneva, Switzerland. *Cancer* 1977;**39**(3):1289-94.

Higgins 2003

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

Howell 1997

Howell A, Dowsett M. Recent advances in endocrine therapy of breast cancer. *BMJ* 1997;**315**:863-6.

Ibrahim 1995

Ibrahim NK, Buzdar AU. Aromatase inhibitors: current status. *American Journal of Clinical Oncology* 1995;**18**(5):407-17.

Mauri 2006

Mauri D, Pavlidis N, Polyzos NP, Ioannidis JP. Survival with aromatase inhibitors and inactivators versus standard hormonal therapy in advanced breast cancer: meta-analysis. *Journal of the National Cancer Institute* 2006;**98**(18):1285-91.

Miller 1996a

Miller WR. Aromatase inhibitors. *Endocrine-related Cancer* 1996;**3**:65-79.

Parmar 1995

Parmar MKB, Machin D. Survival analysis: A Practical Approach. John Wiley, 1995.

Roseman 1997

Roseman BJ, Buzdar AU, Singletary SE. Use of aromatase inhibitors in postmenopausal women with advanced breast cancer. *Journal of Surgical Oncology* 1997;**66**:215-20.

References to other published versions of this review

Cochrane 2007

Gibson LJ, Dawson CK, Lawrence DH, Bliss JM. Aromatase inhibitors for treatment of advanced breast cancer in postmenopausal women. *Cochrane Database of Systematic Reviews* 2007, Issue 1.

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Alonso-Munoz 1988

Study characteristics	
Methods	Spain, multicentre, N = 105, Dec 1982 - Dec 1985 Three arm trial (only two arms included in review N = 70) Randomisation method not given

Alonso-Munoz 1988 *(Continued)*

	Baseline characteristics balanced
Participants	Age range 37 - 75y Proven metastatic breast cancer, measurable disease sites No previous endocrine therapy
Interventions	AG (500mg for 2w, then 100mg) versus TAM 40mg versus AG + TAM 40mg Numbers in each treatment arm: 35 versus 35 versus 35 (AG+TAM arm data excluded from review N = 35) Assessable patients (two included arms): 31 versus 34 Patients evaluable for toxicity (two included arms): 33 versus 34
Outcomes	Toxicity, TTP, response rate Not survival
Notes	11 not evaluable (4 AG, 6 TAM + AG, 1 TAM) due to: 4 died within 6w, 1 discontinued treatment, 5 toxicity, 1 lost to FU FU duration not given TTP not given by treatment arm

Risk of bias

Bias	Authors' judgement	Support for judgement
Blinding (performance bias and detection bias) All outcomes	High risk	

Bezwoda 1998

Study characteristics

Methods	South Africa, multicentre, N = 96 Double-blind, double-dummy Balanced block stratification by centre Baseline slight imbalance in ER status: 28% versus 20% ER+
Participants	Age range 44 - 82y Measurable or evaluable metastatic breast cancer Prior TAM treatment No previous treatment with AI ECOG perf status < 3
Interventions	Fadrozole 2mg versus MA 160mg Numbers in each treatment arm: 46 versus 50 Assessable patients: 46 versus 50 Patients evaluable for toxicity: 46 versus 50 Treatment until progression or for 1y; median duration 20w
Outcomes	Primary - response rate, TTP, TTF, survival Secondary - QOL, performance status, pain assessment
Notes	FU to relapse or death Median FU not stated Intention-to-treat analysis Subsidiary analysis on a per protocol basis (41 versus 43)

Bezwoda 1998 *(Continued)*

	7 major protocol violations, 2 refusals, 1 early death, 1 lost to FU (numbers not consistent)

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	balanced block stratification by centre
Allocation concealment (selection bias)	Low risk	adequate
Blinding (performance bias and detection bias) All outcomes	Low risk	double-blind, double-dummy

Bonneterre 2001

Study characteristics

Methods	International, multicentre trial, combined results of two trials Feb 1996 - July 1998 97 sites in US and Canada, N = 353 83 sites in Europe, Australia, New Zealand, South Africa, South Amercia, N = 668 Total randomised = 1021 Double-blind, double-dummy Baseline characteristics well-balanced
Participants	Age range 30 - 92y Advanced or metastatic breast cancer
Interventions	Anastrozole 1mg versus TAM 20mg Numbers in each treatment arm: 171 versus 182 (N America) and 340 versus 328 (rest of world) Assessable patients: 511 versus 510 Patients evaluable for toxicity: 506 versus 511 Treatment continued until disease progression
Outcomes	Primary - objective response, TTP, tolerability Secondary - TTF, survival
Notes	FU to progression and death Median FU not known Number of dropouts not given

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment (selection bias)	Low risk	adequate
Blinding (performance bias and detection bias) All outcomes	Low risk	double-blind, double-dummy

Buzdar 1996a

Study characteristics	
Methods	International, multicentre. 122 centres: 49 in North America, 73 in Europe, Australia, South Africa, Double-blind anastrozole, open megestrol acetate Randomisation method - blocks of 6 (Europe), blocks of 3 (N America), parallel groups Two trials combined (N = 764): North America (N = 346) and Europe, Australia, South Africa (N = 378) Three-arm trial (only two arms included in review N = 516) Baseline: apparent imbalance in one treatment group (believed to be artefact)
Participants	Age range 29 - 97y Advanced breast cancer Progressed on anti-oestrogen for advanced disease or progressed on or during adjuvant TAM WHO perf status < 3
Interventions	anastrozole 1mg versus anastrozole 10mg versus MA 160mg Numbers in each treatment arm: 263 versus 248 versus 253 (anastrozole 10mg arm excluded from review N = 248) Assessable patients (two included arms): 263 versus 253 Patients evaluable for toxicity (two included arms): 262 versus 253 Treatment continued until disease progression or withdrawal from treatment for other reasons
Outcomes	Primary - TTP, tumour response, tolerability Secondary - TTF, response duration, survival Clinical assessment every 4w until week 24, every 12w until week 48, then every 3m until progression
Notes	FU median duration 6m 3 no treatment, 1 wrong treatment, 8 lost to FU Intention-to-treat analysis

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	blocks of 6 (Europe), blocks of 3 (N America), parallel groups
Allocation concealment (selection bias)	Low risk	adequate
Blinding (performance bias and detection bias) All outcomes	Low risk	double-blind anastrozole, open megestrol acetate

Buzdar 1996b

Study characteristics	
Methods	Protocol 03 Multicentre, 47 sites, N = 380 Feb 1989 - Dec 1991 Double-blind, parallel, controlled equivalence Randomisation method not specified
Participants	Age range 35 - 92y

Buzdar 1996b *(Continued)*

	Metastatic breast cancer At least one prior hormonal treatment for metastic disease more than 3m previously Prior AI use an exclusion Performance status < 3
Interventions	Fadrozole 2mg versus MA 160mg Numbers in each treatment arm: 196 versus 184 Drug code broken 18m after end of enrolment Assessable patients: 195 versus 184 Patients evaluable for toxicity: 196 versus 184 Treatment continued until disease progression
Outcomes	Objective response rate, TTP, survival, toxicity, duration of response, survival, QOL
Notes	Published together with protocol 06 (Buzdar 1996c) FU until progression Intention-to-treat analysis N = 379 1 patient excluded but included in safety and tolerability

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	randomisation method not specified
Allocation concealment (selection bias)	Low risk	adequate
Blinding (performance bias and detection bias) All outcomes	Low risk	double-blind, parallel, controlled equivalence

Buzdar 1996c

Study characteristics

Methods	Protocol 06 Multicentre, 55 sites, N = 303 Oct 1989 - Aug 1992 Double-blind, parallel, controlled equivalence Randomisation method not specified
Participants	Age range 36 - 92y Metastatic breast cancer At least one prior hormonal treatment for metastic disease more than 3m previously Prior AI use an exclusion Performance status<3
Interventions	Fadrozole 2mg versus MA 160mg Numbers in each treatment arm: 152 versus 151 Assessable patients: 150 versus 148 Patients evaluable for toxicity:152 versus 151 Drug code broken 18m after end of enrolment Treatment continued until disease progression
Outcomes	Primary - overall tumour response (TTP, TTF, survival)

Buzdar 1996c *(Continued)*

	Other - earliest diagnosis of PD, tolerability, safety, QOL
Notes	Published together with protocol 03 (Buzdar 1996b) FU: 33m for tumour response/safety (median 5.5m) 45m for survival (median 18 to 20m) Intention-to-treat analysis N = 298 Not designed or powered to detect differences in survival

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	randomisation method not specified
Allocation concealment (selection bias)	Low risk	adequate
Blinding (performance bias and detection bias) All outcomes	Low risk	double-blind, parallel, controlled equivalence

Buzdar 2001

Study characteristics

Methods	International, multicentre, 120 sites in US, Canada, Europe, N = 602 Three-arm trial (only two arms included in review N = 400) Double-blind, double dummy, phase III Randomisation by country w/o stratification by centre Enrolment over 30 months Baseline characteristics no imbalance
Participants	Age range not given Locally advanced/locoregionally recurrent/metastatic breast cancer At least one measurable/assessable lesion Relapsed or progressed while on anti-oestrogen or relapsed within 12m of stopping antioestrogen Chemotherapy for advanced disease allowed KPF >=50%
Interventions	Letrozole 2mg versus letrozole 10mg versus MA 160mg Numbers in each treatment arm: 202 versus 199 versus 201 (letrozole 2mg arm excluded from review N = 202) Assessable patients: 182 versus 180 Patients evaluable for toxicity: 199 versus 201 Treatment continued until disease progression or withdrawal for other reason
Outcomes	Primary - tumour response Secondary - TTF, TTP, survival, QOL
Notes	FU period 48m after the first visit of the last patient randomised Intention-to-treat analysis 23 ineligible and excluded from tumour analyses

Risk of bias

Buzdar 2001 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	randomisation by country w/o stratification by centre
Allocation concealment (selection bias)	Low risk	adequate
Blinding (performance bias and detection bias) All outcomes	Low risk	double-blind, double-dummy

Canney 1988

Study characteristics	
Methods	UK, number of centres not given, N = 218 Randomised without stratification, performed centrally by phone over 24m
Participants	Median age 64y Actively progressive disease Received hormonal therapy with tamoxifen Received no anticancer therapy within preceding 4w
Interventions	AG (250mg for 2w, increased to 500mg if not toxic effect plus 40mg HC) versus high dose MPA 1000mg Numbers in each treatment arm: 106 versus 112 Patients evaluable for toxicity: 106 versus 112
Outcomes	Duration of response, survival, time to response
Notes	FU duration: minumum 9m, median 55w for AG, 57w MPA 7 patients either violated protocol or did not meet entry criteria but included in analyses Crossover on failure No variation between groups in known prognostic variables

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment (selection bias)	Low risk	performed centrally by phone

Chia 2008

Study characteristics	
Methods	International, multicentre, 138 centres, N = 693 Aug 2003 - Nov 2005 Double-blind, double-dummy, phase III Trial acronym = EFECT Randomisation method not given Baseline characteristics well balanced except for ER+/PR+ (56.4% versus 67.5%)

Chia 2008 *(Continued)*

Participants	Age range 32 -91y Locally advanced or metastatic disease Disease progression after prior non-steroidal AI treatment At least one measurable or assessable lesion ER+/PR+ WHO perf status < 3
Interventions	Exemestane 25mg versus fulvestrant 500mg on day 0, 250mg on days 14 and 28, followed by 250mg every four weeks Numbers in each treatment arm: 342 versus 351 Assessable patients: 270 versus 270 Treatment continued until disease progression
Outcomes	Primary - TTP Secondary - objective response, CB, response duration, TTF, overall survival, tolerability, QOL
Notes	FU until death Intention-to-treat analysis 90% power to detect HR≥1.31

Risk of bias

Bias	Authors' judgement	Support for judgement
Blinding (performance bias and detection bias) All outcomes	Low risk	double-blind, double-dummy

Dombernowsky 1998

Study characteristics

Methods	International, multicentre, Mar 1993 - Sep 1994 10 countries, 91 sites, N = 551 Three-arm trial (only two arms included in review N=363) Double-blind, randomisation stratified by country; computer-generated permuted blocks of size 6 or 3, 1:1:1 allocation Baseline characteristics balanced
Participants	Advanced/locoregionally recurrent/metastatic breast cancer Measurable/assessable disease Failure to respond to previous antioestrogen WHO perf status < 3
Interventions	Letrozole 0.5mg versus letrozole 2.5mg versus MA 160mg Numbers in each treatment arm: 188 versus 174 versus 189 (letrozole 0.5mg arm excluded from review N = 188) Assessable patients: 153 versus 166 Patients evaluable for toxicity: 174 versus 189
Outcomes	Primary - overall tumour response (TTP, TTF, survival) Other - earliest diagnosis of PD, tolerability, safety
Notes	FU: 33m for tumour response/safety (median 5.5m) 45m for survival (median 18 to 20m) Intention-to-treat analysis

Dombernowsky 1998 *(Continued)*

	Not designed or powered to detect differences in survival as significant

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	stratified by country; computer-generated permuted blocks of size 6 or 3, 1:1:1 allocation
Allocation concealment (selection bias)	Low risk	adequate
Blinding (performance bias and detection bias) All outcomes	Low risk	double-blind

Falkson 1996

Study characteristics

Methods	South Africa, single centre, N = 80 Sep 1991 - Dec 1994 Randomisation method not given Baseline: difference of 10y in median age of patients in arm 1 versus arm 2
Participants	Age range 43 - 90y Progressive, inoperable, recurrent or metastatic breast cancer No prior treatment for advanced disease ECOG perf status < 3
Interventions	Fadrozole 2mg versus TAM 20mg Numbers in each treatment arm: 40 versus 40 Assessable patients: 36 versus 38 Patients evaluable for toxicity: 40 versus 40 Minimum treatment 8w
Outcomes	Survival, TTF, duration of overall response, toxicity, objective response rates
Notes	FU 14 to 1122d, median FU 153d Intention-to-treat analysis 2 ineligible, 1 lost to FU 74 patients evaluable

Freue 2000

Study characteristics

Methods	International, multicentre, 9 countries, 78 centres, N = 547 Aug 1991 - Mar 1995 Computer-generated random allocation w/o stratification Open study No difference in baseline characteristics

Freue 2000 *(Continued)*

Participants	Age range not given Advanced disease Measurable disease ER/PR positive or unknown WHO perf status < 3 Only TAM as 1st line endocrine therapy
Interventions	Formestane 250mg im every 2w versus MA 160mg Numbers in each treatment arm: 276 versus 271 Assessable patients: 242 versus 237 Numbers for safety analysis: 276 versus 271 Treatment duration 12m
Outcomes	TTF, TTP, overall survival, overall response
Notes	FU until death Median FU not given 90% power to detect 33% difference in median TTF Intention-to-treat analysis Ineligible/non-evaluable: 34 versus 34 Non-cancer deaths: 2 versus 4 Discontinued for AE: 3 versus 13

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	computer-generated random allocation without stratification
Allocation concealment (selection bias)	Low risk	adequate
Blinding (performance bias and detection bias) All outcomes	High risk	open

Gale 1994

Study characteristics

Methods	ECOG trial, multicentre, US, N = 249 1977 - 1983 Stratified randomly permuted blocks of four Baseline characteristics relatively evenly balanced One institution had 60% versus 4% response rates
Participants	Age range not given Progressive, recurrent, metastatic breast cancer Measurable disease ECOG perf status < 4 No previous treatment with AG or TAM
Interventions	AG 250mg qid versus TAM 20mg Numbers in each treatment arm: 122 versus 119 Assessable patients: 108 versus 108

Gale 1994 *(Continued)*

Outcomes	Tumour response, TTF, overall survival
Notes	Initial trial design changed in May 1979 (adrenalectomy treatment arm discontinued) Crossover on progression Crossover results not included Intention-to-treat analysis Adrenalectomy patients (N = 8) were excluded

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	stratified randomly permuted blocks of four
Allocation concealment (selection bias)	Low risk	adequate

Garcia-Giralt 1992

Study characteristics

Methods	France, multicentre, N = 257 No randomisation details
Participants	Age range 36 - 91y Histologically confirmed metastatic breast cancer ER+/PR+ Initial response to TAM before relapse
Interventions	AG 500mg + HC versus MPA 1000mg Numbers in each treatment arm: 131 versus 119 Assessable patients: 124 versus 112 Second-line therapy after TAM
Outcomes	Tumour response, TTP, new metastases, AEs
Notes	Median FU not known Treatment until progression Crossover on progression 6 lost to FU, 1 man

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment (selection bias)	Unclear risk	not used

Gershanovich 1998

Study characteristics

Gershanovich 1998 *(Continued)*

Methods	International, multicentre, 11 countries, 86 sites N = 555 Three-arm trial (only two arms included in review N = 363) Open-label 1:1:1 Baseline no major differences
Participants	Median age letrozole 2.5mg 66y, letrozole 0.5 mg 64y, AG 65y Advanced or metastatic breast cancer Measurable/evaluable advanced disease WHO perf status < 3
Interventions	Letrozole 2.5mg versus letrozole 0.5mg versus AG 500 mg Numbers in each treatment arm: 185 versus 192 versus 178 (letrozole 2.5mg arm excluded from review N = 192) Assessable patients: 173 versus 162
Outcomes	Response, TTP, TTF, survival, tolerability and safety, overall survival
Notes	FU duration median > 20m 44 not assessable, counted as non-responders in the analysis Median duration of treatment 5m Modified intention-to-treat population ie enrolled and received trial medication

Risk of bias

Bias	Authors' judgement	Support for judgement
Blinding (performance bias and detection bias) All outcomes	High risk	

Goss 1999

Study characteristics

Methods	Nov 1991 - Dec 1995 Multicentre, 29 sites in Canada and 38 in US, N = 452 Open-label, stratified by disease status Baseline characteristics comparable
Participants	Age range 39 - 90y Advanced breast cancer, histologically confirmed Progressed after tamoxifen treatment
Interventions	Vorozole 2.5mg versus MA 160mg Numbers in each treatment arm: 225 versus 227 Assessable patients: 190 versus 185 Patients evaluable for toxicity: 195 versus 198 2nd line treatment after tamoxifen
Outcomes	Primary - response rate Secondary - TTP, survival, duration of response, safety subjective symptoms, QOL
Notes	Median FU 11.6m (vorozole), 9.9m (MA) 1 withdrawn before treatment 4 ineligible, 18 adverse events, 1 lost to FU, 18 other

Goss 1999 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	stratified by disease status
Allocation concealment (selection bias)	Low risk	adequate
Blinding (performance bias and detection bias) All outcomes	High risk	open-label

Goss 2007

Study characteristics	
Methods	Multinational, multicentre, 60 centres in US, Canada, Russia, Ukraine, N = 865 Randomised, double-blind, active control, phase III Randomisation in blocks of four, stratified by centre. Performed centrally, site notified by fax Treatment code unblinded after database lock Baseline characteristics well balanced
Participants	Median age letrozole 63y atamestane 65y Locally recurrent/advanced/ metastatic disease Measurable disease No AI or antioestrogen/SERM treatment in previous 12m ECOG perf status < 3
Interventions	Letrozole 2.5mg versus atamestane 500mg + toremifene 60mg Numbers in each treatment arm: 431 versus 434 Assessable patients: 297 versus 298
Outcomes	Primary - TTP Secondary - overall survival, TTF, tumour response, toxicity
Notes	FU to death Intention-to-treat analysis Treatment continued until disease progression or withdrawal for other reasons 80% power to detect a 24% increase in TTP

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	blocks of four, stratified by centre
Allocation concealment (selection bias)	Low risk	performed centrally, site notified by fax
Blinding (performance bias and detection bias) All outcomes	Low risk	double-blind

Ingle 1986

Study characteristics	
Methods	US, number of centres not known, N = 102 Randomised using Pocock-Simon approach to adaptive randomisation, stratified
Participants	Age range 38 - 83y Progressive metastatic disease Measurable or evaluable lesion ECOG perf status < 4 No prior therapy with either AG or TAM
Interventions	TAM 20mg versus TAM (20mg) + AG (500mg for 2 weeks then 1000mg) + HC (100mg daily for 2 weeks then 40mg) Numbers in each treatment arm: 49 versus 51 Assessable patients:49 versus 51 Patients evaluable for toxicity: 48 versus 46
Outcomes	Objective response, TTP, survival, toxicity
Notes	No data on duration of FU Target accrual = 160 but terminated early due to excess toxicity on the TAM + AG + HC arm 2 patients ineligible

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	used Pocock-Simon approach to adaptive randomisation, stratified

Kaufmann 2000

Study characteristics	
Methods	International, multicentre, Oct 1995 - May 1998 19 countries, 144 centres N = 769 Double-blind, parallel-group, phase III Baseline characteristics comparable
Participants	Age range 30 - 91y Advanced breast cancer Progressed or relapsed during tamoxifen treatment
Interventions	Exemestane 25mg versus MA 160mg Numbers in each treatment arm: 366 versus 403 Assessable patients: 337 versus 366 Patients evaluable for toxicity: 358 versus 400
Outcomes	Objective response, TTP, TTF, survival, tumour response, duration of tumour control, tumour related signs and symptoms, QOL, tolerability
Notes	FU median duration 48.9w 6 randomised but not treated 66 not evaluable for tumour response

Kaufmann 2000 *(Continued)*

	Intention-to-treat analysis

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment (selection bias)	Low risk	adequate
Blinding (performance bias and detection bias) All outcomes	Low risk	double-blind

Kleeberg 1997

Study characteristics

Methods	International, multicentre, 27 Jun - 1 Dec 1995 18 centres, Europe and South Africa, N = 60 open-label, parallel-group, comparative Baseline good balance re age, weight, prior tamoxifen treatment
Participants	Age range 40 - 84y Advanced breast cancer Measurable or evaluable disease
Interventions	Anastrozole 1mg oral per day versus formestane 250mg im every 2w Numbers in each treatment arm: 29 versus 31 Assessable patients: 29 versus 31 Treatment until disease progression
Outcomes	Primary - oestradiol suppression and tolerability Secondary - response rates, TTP, adverse events, blood oestrone sulphate, patient and doctor perception of treatment
Notes	No details re randomisation exclusions or FU Not powered to detect clinically significant difference in oestrogen suppression between the two arms

Risk of bias

Bias	Authors' judgement	Support for judgement
Blinding (performance bias and detection bias) All outcomes	High risk	open-label

Leitzel 1995

Study characteristics

Methods	Location and date of trial not given Multicentre, N = 300 Double-blind, double-dummy, parallel

Leitzel 1995 *(Continued)*

	Randomisation method not given
Participants	Age range 18 - 85y Metastatic breast cancer ECOG < 3
Interventions	Fadrozole 2mg versus MA 160mg Numbers in each treatment arm not given Duration of intervention not given Second-line treatment
Outcomes	Tumour response, progression, c-erbB-2 antigen in serum
Notes	FU until death Results not given by treatment group Survival was not given by treatment group although it was measured

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment (selection bias)	Low risk	adequate
Blinding (performance bias and detection bias) All outcomes	Low risk	double-blind, double-dummy

Lundgren 1989

Study characteristics

Methods	Norway, multicentre, N = 176 Randomisation without stratification, details not given Baseline characteristics well balanced for most important prognostic variables, except main metastatic site
Participants	Mean age 62.0y versus 62.7y Advanced breast cancer Evaluable disease Previous treatment with TAM KPS >50
Interventions	AG 250mg bid for 2w then 250mg qid versus MA 160mg Numbers in each treatment arm: 86 versus 90 Assessable patients: 76 versus 74 Second-line treatment
Outcomes	Response rate, reponse duration, survival, toxicity
Notes	Intention-to-treat analysis Excluded patients: 10 protocol violations/patient refusal; 12 early deaths; 4 adverse events

Risk of bias

Bias	Authors' judgement	Support for judgement

Lundgren 1989 *(Continued)*

Random sequence generation (selection bias)	Unclear risk	randomisation without stratification, details not given
Allocation concealment (selection bias)	Unclear risk	not used

Mauriac 2003

Study characteristics	
Methods	Data were combined and published from two trials 0020 and 0021 (May 1997 - September 1999) Trial 0020: multicentre, phase III, open, parallel group Europe, Australia and South Africa, 83 centres, N = 451 Trial 0021: multicentre, phase III, double-blind, double-dummy, parallel group North America, N = 400 Combined data from both trials included in review N = 851
Participants	Age range 33 - 89y Locally advanced or metastatic breast cancer Progressed during adjuvant endocrine therapy or first-line therapy for advanced disease WHO performance status < 3
Interventions	Anastrozole 1mg versus fulvestrant 250mg/month im Trial 0020: numbers in each treatment arm: 222 versus 229 Trial 0021: numbers in each treatment arm: 206 versus 194 Combined trials (included in review): numbers in each treatment arm: 423 versus 428 Assessable patients: 423 versus 428 Patients evaluable for toxicity: 423 versus 423 Continued until objective disease progression or other events required withdrawal
Outcomes	TTP, objective response, tolerability, QOL
Notes	Median FU 15.1m (combined data) Intention-to-treat analysis Additional to protocol: non-inferiority of fulvestrant with anastrozole was carried out retrospectively

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Allocation concealment (selection bias)	Low risk	adequate
Blinding (performance bias and detection bias) All outcomes	Low risk	Trial 0020 open Trial 0021 double-blind

Mercer 1993

Study characteristics	
Methods	UK, query single-centre, Jan 1987 - Dec 1990, N = 61 No information regarding randomisation

Mercer 1993 *(Continued)*

	Groups well matched but after exclusions numbers small
Participants	Eligibility >50y Age range 45 - 86y Advanced breast cancer Progressive disease on tamoxifen (adjuvant or treatment)
Interventions	Low dose AG 125mg versus HC 20mg Number in each treatment arm: 28 versus 33 Assessable patients: 27 versus 29
Outcomes	Tumour response, TTF, side-effects, overall survival
Notes	FU details not given 5 patients excluded

Milla-Santos 2003

Study characteristics	
Methods	Spain, single-centre, N = 238, May 1997 - Dec 1999 Randomisation following Meinert's methodology. Baseline characteristics comparable
Participants	Age range 55 - 77y Histologically confirmed advanced breast cancer, measurable disease sites No previous endocrine therapy ECOG<3
Interventions	Anastrozole 1mg versus TAM 40mg Numbers in each treatment arm: 121 versus 117 Assessable patients: 121 versus 117
Outcomes	Primary - response rates, clinical benefit, TTP in patients achieving a CB, overall survival, toxicity
Notes	FU to 35m intention-to-treat analysis All patients evaluable Analysis cutoff 1 April 2001

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	randomisation following Meinert's methodology

Mourisden 2001

Study characteristics	
Methods	International, multicentre, Nov 1996 - Jan 1999 29 countries, 201 sites, N = 939 Double-blind, double-dummy, parallel group

Mourisden 2001 *(Continued)*

	Baseline characteristics well balanced
Participants	Age range 31 - 96y Locally advanced/locoregionally recurrent/metastatic breast cancer which is measurable/assessable Previous chemotherapy allowed for advanced disease WHO perf status < 3
Interventions	Letrozole 2.5mg versus TAM 20mg Numbers in each treatment arm: 453 versus 454 Assessable patients: 421 versus 423 Patients evaluable for toxicity: 455 versus 455 Treatment continued until disease progression
Outcomes	Primary - TTP Secondary - tumour response rate, TTF, ORR, survival, tolerability, KPS
Notes	FU median 32m Intention-to-treat analysis 907 analysed, 32 excluded Analysis cutoff March 2000 Survival not reported 729 discontinued treatment of which 391 'crossed over'

Risk of bias

Bias	Authors' judgement	Support for judgement
Allocation concealment (selection bias)	Low risk	adequate
Blinding (performance bias and detection bias) All outcomes	Low risk	double-blind, double-dummy

Paridaens 2003

Study characteristics

Methods	International, multicentre, October 96 - May 99 13 centres in 6 countries, N = 122 Open-label phase II, randomised centrally using minimisation by EORTC, stratified by centre, adjuvant TAM, CT for metastatic disease, dominant disease site The trial was designed as a randomised phase II trial not to enable comparison of the efficacy of the two drugs but to establish a 'go, no-go' rule for exemestane activity and safety before a formal randomised phase III trial. Patients randomised into the phase II trial will be incorporated into the phase III trial
Participants	Age range 37 - 87y Measurable metastatic or locally recurrent inoperable breast cancer No prior hormone therapy for metastatic disease ECOG perf status < 3
Interventions	Exemestane 25mg versus TAM 20mg Numbers in each treatment arm: 62 versus 60 Intention-to-treat analysis: 61 versus 59 Toxicity data: 62 versus 59 Assessable patients: 56 versus 57

Paridaens 2003 *(Continued)*

	Patients evaluable for toxicity: 62 versus 59 Treatment continued until disease progression
Outcomes	Response rates Stop-go for phase III Phase II therefore inadequate power, no statistical comparison of efficacy of endpoints between the two treatments were planned or performed
Notes	FU details 2 patients (1 exemestane, 1 TAM) ineligible as not having metastatic breast cancer, 7 additional (5 exemestane, 2 TAM) not evaluable for response, 1 lost to FU Phase II patients to be included in phase III trial Intention-to-treat analysis

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	stratified by centre, randomised centrally by EORTC using minimisation
Allocation concealment (selection bias)	Low risk	adequate
Blinding (performance bias and detection bias) All outcomes	High risk	open-label

Perez Carrion 1994

Study characteristics

Methods	International, multicentre, May 1988 - December 1990, N = 409 Open study, equivalence trial Baseline characteristics well matched
Participants	Age range 38 - 87y WHO perf status < 3
Interventions	Formestane 250mg im versus TAM 30mg Numbers in each treatment arm: 203 versus 206 Assessable patients: 173 versus 175
Outcomes	Response, survival, TTP, TTF, tolerability
Notes	FU details not reported 61 patients not evaluable, 10 lost to FU, 3 refusals Intention-to-treat analysis Trial closed early due to changes in clinical practice, ie increasing use of TAM in the adjuvant setting

Risk of bias

Bias	Authors' judgement	Support for judgement
Blinding (performance bias and detection bias)	High risk	open

Perez Carrion 1994 *(Continued)*

All outcomes

Powles 1984

Study characteristics		
Methods	Sept 1979 - June 1983 UK, single-centre, N = 222 Previously determined allocation list unknown to clinician. Baseline characteristics mean age marginally greater for TAM patients	
Participants	Patients with disseminated breast cancer who had not previously received TAM, AG, or danazol No endocrine or chemotherapy within 6w	
Interventions	TAM 20mg versus TAM 20mg + AG 750mg + danazol 300mg + HC 40mg Number on each treatment arm: 111 versus 111 Assessable patients: 99 versus 99 Patients evaluable for toxicity: 111 versus 111 Treatment continued until 3m assessment (unless rapid development of tumour in meantime) otherwise stopped when evidence of tumour progression arose either through failure to respond or because of relapse after response or stabilisation of disease	
Outcomes	Tumour response	
Notes	FU duration not reported	
Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Allocation concealment (selection bias)	Low risk	previously determined allocation list unknown to clinician

Rose 1986

Study characteristics	
Methods	Denmark, multicentre, June 1979 - Sept 1988, 4 centres N = 313 Three-arm trial (only two arms included in review N = 215) Randomised by centre, non-stratified, stochastic array of numbers, closed envelope system Baseline characteristics well balanced
Participants	Age > 65y, age range 66 - 84y First recurrence of metastatic breast cancer Progressive disease with measurable and/or evaluable lesions Performance status < 4
Interventions	TAM 30mg versus TAM 30mg + AG 250mg qid + HC 60mg v TAM 30mg + fluoxymesterone 20mg Numbers in each treatment arm: 108 versus 107 versus 98 (TAM + fluoxymesterone excluded from review N = 98) Assessable patients: 83 versus 94 Patients evaluable for toxicity: 87 versus 97 Treatment until progression (minimum 12 weeks)

Rose 1986 *(Continued)*

Outcomes	TTF, TTP, survival, toxicity
Notes	FU duration not reported 34 ineligible 21 not evaluable 9 lost to FU 258 fully evaluable

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	randomised by centre, non-stratified, stochastic array of numbers
Allocation concealment (selection bias)	Low risk	closed envelope system

Rose 2003

Study characteristics

Methods	International phase IIIb/IV, 19 countries, multicentre, 112 sites, N = 713 Dec 1997 - Nov 1999 Open, random assignation stratified by centre via predetermined randomisation list Baseline characteristics well balanced
Participants	Age range 27 - 92y Advanced or metastatic breast cancer with measurable and/or evaluable disease Histologically/cytologically confirmed Previous treatment with antioestrogen WHO performance status 0-2
Interventions	Letrozole 2.5mg versus anastrozole 1mg Numbers in each treatment arm: 356 versus 357 Assessable patients: 299 versus 304
Outcomes	Primary - TTP Secondary- objective response, duration of response, rate and duration of overall clinical benefit, overall survival, general safety
Notes	FU duration not reported

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	random assignation stratified by centre
Allocation concealment (selection bias)	Low risk	predetermined randomisation list
Blinding (performance bias and detection bias)	High risk	open-label

Rose 2003 *(Continued)*

All outcomes

Russell 1997

Study characteristics	
Methods	May 1984 - November 1990, Phase III, N = 288 Three-arm trial (only two arms included in review N = 155) No stratification Treatment arms reasonably well balanced
Participants	Age range 33 - 92y Progressive metastatic disease Measurable or evaluable lesion Patients had received TAM in advanced setting No prior MA or AG
Interventions	MA 160mg versus AG (500mg for 2w then 1000mg) + HC (100mg for 2w then 40mg) versus MA 160mg + AG (500mg for 2w then 1000mg) + hydrocortisone Numbers in each treatment arm: 75 versus 80 versus 80 (MA 160mg + AG (500mg for 2w then 1000mg) + hydrocortisone arm data excluded from review N = 80) Assessable patients: 42 versus 32 Patients evaluable for toxicity: 88 versus 89
Outcomes	Response, TTF, survival, toxicity
Notes	FU median duration amongst those still alive = 5.2y (213 had died) 53 ineligible (38 re misunderstanding re prior TAM use,7 due to life threatening visceral involvement, 3 with less than 6 months of TAM, 2 ER -, 1 prior hormonal therapy other than TAM, 1 no confirmed disease sites) Patients on MA or AG alone were crossed over after progression

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	no stratification

Samonis 1994

Study characteristics	
Methods	Greece, single-centre, N = 85 trial duration 2.5y Three-arm trial (only two arms included in the review N = 57) Stratified randomisation - statified into four groups by previous adjuvant treatment Table of baseline characteristics
Participants	Age range 50 - 73y Metastatic breast cancer Measurable disease No previous treatment with AG or MPA KPS > 70%

Samonis 1994 *(Continued)*

Interventions	AG (250mg for 3d, then to 1000mg) versus MPA (500mg for 1m then twice weekly) versus AG + MPA Numbers in each treatment arm:28 versus 29 versus 28 (AG + MPA data excluded from review) Assessable patients (two included arms): 26 versus 27
Outcomes	Response to treatment, toxicity
Notes	FU duration not given Excluded patients: 1 accidental death, 4 lost to FU

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	statified into four groups by previous adjuvant treatment

Schmid 2001

Study characteristics

Methods	International, multicentre, N = 171 Three-arm trial (only two arms included in review N = 112) Double-blind
Participants	Mean age 64.5 Advanced breast cancer with bone metastases
Interventions	Letrozole 2.5mg versus letrozole 0.5mg versus MA 160mg Number in each treatment arm: 52 versus 59 versus 60 (letrozole 0.5mg arm excluded from review N = 59) Assessable patients: 48 versus 53
Outcomes	Objective response, clinical benefit, TTP, survival
Notes	Publication only available as abstract but sufficient data to include

Risk of bias

Bias	Authors' judgement	Support for judgement
Blinding (performance bias and detection bias) All outcomes	Low risk	double-blind

Thuerlimann 1996

Study characteristics

Methods	Switzerland, Phase III multicentre, 7 sites, N = 221 June 1988 - Dec 1994 Phone randomisation, stratified, minimisation not double blind Baseline: prognostic factors well balanced apart from metastatic site

Thuerlimann 1996 *(Continued)*

Participants	Age range 39 - 87y Measurable/evaluable advanced breast cancer Indication for hormone treatment ECOG < 2
Interventions	Fadrozole 2mg versus TAM 20mg Numbers in each treatment arm: 111 versus 110 Eligible patients: 105 versus 107 Assessable patients: 103 versus 106 Patients evaluable for toxicity: 104 versus 107 First-line treatment Treatment until progression
Outcomes	TTF, response rate, toxicity, overall survival, TTP, subjective benefit (not reported), duration of response
Notes	FU 7½ y Eligible patients: 212 9 ineligible(6 fadrozole, 3 TAM) 12 withdrawals Crossover only after failure so not analysed Analysis on data to Dec 1995, median FU of survivors 3y

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	minimisation, stratified
Allocation concealment (selection bias)	Low risk	phone randomisation
Blinding (performance bias and detection bias) All outcomes	Unclear risk	not double blind

Thuerlimann 1997

Study characteristics

Methods	Feb 1991 - Jun 1995, N = 179 Stratified, central randomisation Baseline characteristics well balanced (only difference in weight)
Participants	Age range 43 - 87y Advanced breast cancer Histologically and/or cytologically proven with measurable/evaluable disease Failed prior adjuvant and/or palliative tamoxifen treatment ie second-line treatment Prior chemotherapy allowed ECOG perf status < 3
Interventions	Formestane 250mg im (biweekly) versus MA 160mg Numbers in each treatment arm: 91 versus 86 Assessable patients: 90 versus 83 Patients evaluable for toxicity: 90 versus 81

Thuerlimann 1997 *(Continued)*

Outcomes	TTF, toxicity
Notes	FU duration not reported 2 ineligible, 4 dropouts 173 fully evaluable After failure of randomised treatment 75 patients 'crossed over'

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	stratified
Allocation concealment (selection bias)	Low risk	central randomisation

Tominaga 2003

Study characteristics

Methods	Japan, multicentre, 62 sites, N = 157 Double-blind, double-dummy, parallel groups Adaptive dynamic balancing method
Participants	Mean age 59.7y (letrozole) and 61.0y (fadrozole) Advanced disease Measurable or assessable pathological lesions
Interventions	Letrozole 1mg versus fadrozole 2mg Numbers in each treatment arm: 79 versus 78 Assessable patients: 77 versus 77 Minimum 8w treatment Treatment until disease progressed or patient experienced toxicity resulting in discontinuation
Outcomes	ORR, safety of letrozole compared to fadrozole
Notes	FU median 13.3m

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	adaptive dynamic balancing method
Allocation concealment (selection bias)	Low risk	adequate
Blinding (performance bias and detection bias) All outcomes	Low risk	double-blind

KPS - Karnofsky Performance Status
AG - aminoglutethimide

AI - aromatase inhibitor
CB - clinical benefit
ECOG - Eastern Cooperative Oncology Group
EORTC - European Organization for the Research and Treatment of Cancer
ER - oestrogen receptor
FU - follow up
im - intramuscular
mg - milligram
TAM - tamoxifen
MA - megestrol acetate
MPA - medroxy progesterone acetate
HC - hydrocortisone
N - number of patients
ORR - objective response rate
PD - progressive disease
perf status - performance status
qid - four times daily
QOL - quality of life
TTF - time to failure
TTP - time to progression
d - days
w - weeks
m - months
y - years
WHO - World Health Organisation
w/o - without

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Abe 2002	dose comparison of same AI (letrozole)
Bajetta 1994	dose comparison of same AI (formestane)
Bajetta 1997	dose comparison of same AI (exemestane)
Bajetta 1997a	dose comparison of same AI (letrozole)
Bajetta 1999	dose comparison of same AI (letrozole)
Beretta 1990	dose comparison of same AI (letrozole)
Bruning 1989	dose comparison of same AI (aminoglutethimide)
Bruning 1990	dose comparison of same AI (aminoglutethimide)
Castelazo 2004	non-English (Spanish) paper
Cataliotti 2006	comparison of anastrozole versus tamoxifen as neoadjuvant treatment
Dixon 2000	dose-comparison of same AI (anastrozole)
Dowsett 1989	dose-comparison of same AI (formestane)
Dowsett 1990	dose-comparison of same AI (fadrozole)
Dowsett 1994	dose-comparison of same AI (fadrozole)

Study	Reason for exclusion
Dowsett 1995	dose-comparison of same AI (letrozole)
Eiermann 2001	comparison of letrozole versus tamoxifen as pre-operative treatment
Geisler 1996	outcome: aromatase levels and plasma oestrogen levels
Geisler 2002	outcome: aromatase levels and plasma oestrogen levels
Ingle 1997	dose comparison of same AI (letrozole)
Johnston 1994	dose comparison of same AI (vorozole)
Miller 1996b	dose comparison of same AI (fadrozole)
Pronzato 1993	AI (aminoglutethimide) versus same AI plus tamoxifen
Raats 1992	dose comparison of same AI (fadrozole)
Smith 2005	comparison of anastrozole versus tamoxifen as neoadjuvant treatment
Svenstrup 1994	dose comparison of same AI (fadrozole)
Wang 2003	non-English (Chinese) paper

Characteristics of ongoing studies *[ordered by study ID]*

CAAN

Study name	CAAN
Methods	
Participants	Target accrual = 90 postmenopausal women with histologically proven advanced breast cancer
Interventions	Exemestane + celecoxib versus exemestane versus letrozole
Outcomes	Levels of serum lipids and cholesterol
Starting date	February 2002
Contact information	LWC Chow lwcchow@hkucc.hku.hk
Notes	initial report published in 2005

ECOG E4101

Study name	ECOG E4101
Methods	

ECOG E4101 *(Continued)*

Participants	Target accrual = 148 postmenopausal women with HR+ metastatic breast cancer previously treated with up to two chemotherapy regimens and/or one prior endocrine therapy
Interventions	Faslodex + gefitinib versus arimidex + gefitinib
Outcomes	
Starting date	
Contact information	Dr RW Carlson or AstroZeneca
Notes	currently recruiting in the USA

ICR-CTSU Sofea

Study name	Sofea Phase III
Methods	
Participants	Target accrual = 750 women with metastatic disease who have failed after non-steroidal AI
Interventions	Faslodes versus faslodex + anastrozole versus exemestane
Outcomes	
Starting date	March 2004
Contact information	Dr SRD Johnston, Royal Marsden Hospital email: sofea-icrctsu@icr.ac.uk
Notes	Open to recruitment in UK

Paridaens 2003

Study name	Phase III EORTC-10951
Methods	
Participants	Postmenopausal women with metastatic and progressive disease or locally recurrent and inoperable
Interventions	exemestane versus tamoxifen
Outcomes	
Starting date	
Contact information	robert.paridaens@uz.kuleven.ac.be
Notes	phase II to phase III study

HR+ HER positive

DATA AND ANALYSES

Comparison 1. AI versus non-AI

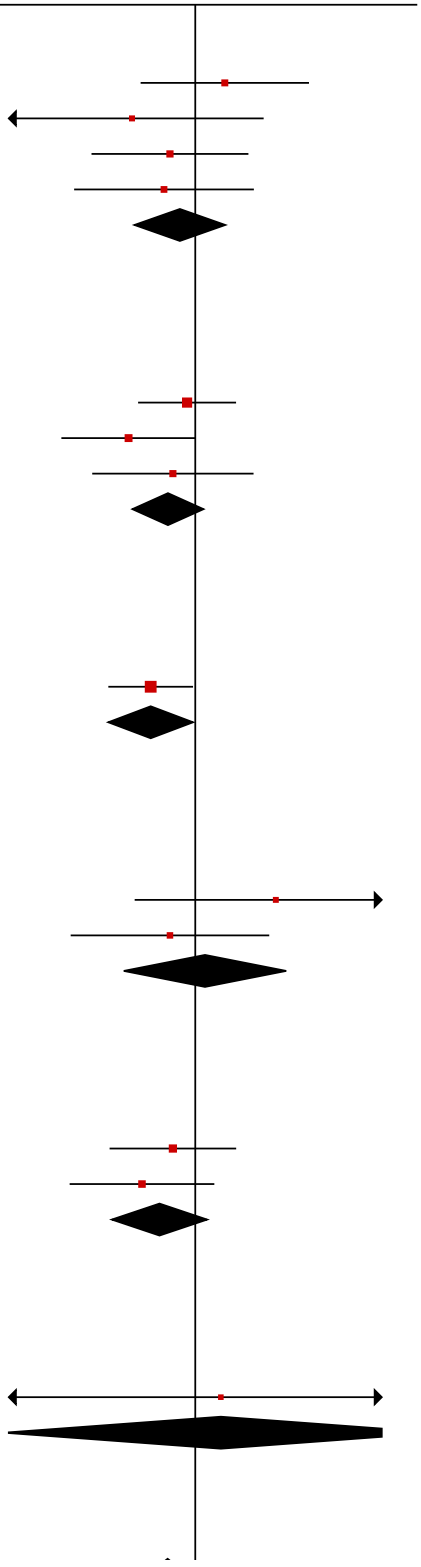
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.1 Overall survival (reported or calculated)	13		HR (IV, Fixed, 95% CI)	0.90 [0.84, 0.97]
1.1.1 aminoglutethimide (any dose)	4		HR (IV, Fixed, 95% CI)	0.94 [0.80, 1.12]
1.1.2 anastrozole 1 mg	3		HR (IV, Fixed, 95% CI)	0.90 [0.79, 1.03]
1.1.3 exemestane 25 mg	1		HR (IV, Fixed, 95% CI)	0.85 [0.72, 0.99]
1.1.4 fadrozole 2 mg	2		HR (IV, Fixed, 95% CI)	1.04 [0.77, 1.40]
1.1.5 letrozole 2.5 mg	2		HR (IV, Fixed, 95% CI)	0.88 [0.73, 1.05]
1.1.6 vorozole 2.5 mg	1		HR (IV, Fixed, 95% CI)	1.10 [0.49, 2.47]
1.2 Progression-free survival (reported or calculated)	11		HR (IV, Random, 95% CI)	0.98 [0.84, 1.13]
1.2.1 aminoglutethimide (any dose)	2		HR (IV, Random, 95% CI)	1.07 [0.73, 1.55]
1.2.2 formestane 250 mg	1		HR (IV, Random, 95% CI)	0.93 [0.68, 1.28]
1.2.3 anastrozole 1 mg	2		HR (IV, Random, 95% CI)	1.05 [0.65, 1.70]
1.2.4 exemestane 25 mg	2		HR (IV, Random, 95% CI)	0.91 [0.72, 1.14]
1.2.5 letrozole 2.5 mg	3		HR (IV, Random, 95% CI)	0.87 [0.68, 1.11]
1.2.6 vorozole 2.5 mg	1		HR (IV, Random, 95% CI)	1.27 [1.04, 1.56]
1.3 Clinical benefit (assessable)	27	8789	Odds Ratio (M-H, Random, 95% CI)	0.87 [0.77, 0.99]
1.3.1 aminoglutethimide (any dose)	9	1292	Odds Ratio (M-H, Random, 95% CI)	0.79 [0.63, 1.00]
1.3.2 formestane 250 mg	2	521	Odds Ratio (M-H, Random, 95% CI)	1.05 [0.59, 1.86]
1.3.3 anastrozole 1 mg	4	2626	Odds Ratio (M-H, Random, 95% CI)	0.74 [0.48, 1.12]
1.3.4 exemestane 25 mg	3	1356	Odds Ratio (M-H, Random, 95% CI)	0.86 [0.63, 1.19]
1.3.5 fadrozole 2 mg	4	982	Odds Ratio (M-H, Random, 95% CI)	1.05 [0.80, 1.38]
1.3.6 letrozole 2.5 mg	4	1637	Odds Ratio (M-H, Random, 95% CI)	0.77 [0.60, 1.00]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.3.7 vorozole 2.5 mg	1	375	Odds Ratio (M-H, Random, 95% CI)	1.35 [0.88, 2.07]
1.4 Objective response (assessable)	31	9595	Odds Ratio (M-H, Random, 95% CI)	0.88 [0.77, 1.01]
1.4.1 aminoglutethimide (any dose)	11	1545	Odds Ratio (M-H, Random, 95% CI)	0.83 [0.63, 1.09]
1.4.2 formestane 250 mg	3	1000	Odds Ratio (M-H, Random, 95% CI)	1.23 [0.92, 1.64]
1.4.3 anastrozole 1 mg	4	2626	Odds Ratio (M-H, Random, 95% CI)	0.95 [0.77, 1.17]
1.4.4 exemestane 25 mg	3	1356	Odds Ratio (M-H, Random, 95% CI)	0.67 [0.33, 1.33]
1.4.5 fadrozole 2 mg	5	1056	Odds Ratio (M-H, Random, 95% CI)	1.18 [0.85, 1.65]
1.4.6 letrozole 2.5 mg	4	1637	Odds Ratio (M-H, Random, 95% CI)	0.65 [0.51, 0.82]
1.4.7 vorozole 2.5 mg	1	375	Odds Ratio (M-H, Random, 95% CI)	0.70 [0.34, 1.42]
1.5 Clinical benefit (randomised)	27	9425	Odds Ratio (M-H, Random, 95% CI)	0.88 [0.78, 0.99]
1.5.1 aminoglutethimide (any dose)	9	1395	Odds Ratio (M-H, Random, 95% CI)	0.81 [0.65, 1.01]
1.5.2 formestane 250 mg	2	586	Odds Ratio (M-H, Random, 95% CI)	1.00 [0.58, 1.70]
1.5.3 anastrozole 1 mg	4	2626	Odds Ratio (M-H, Random, 95% CI)	0.74 [0.48, 1.12]
1.5.4 exemestane 25 mg	3	1584	Odds Ratio (M-H, Random, 95% CI)	0.88 [0.71, 1.11]
1.5.5 fadrozole 2 mg	4	1000	Odds Ratio (M-H, Random, 95% CI)	1.08 [0.82, 1.41]
1.5.6 letrozole 2.5 mg	4	1782	Odds Ratio (M-H, Random, 95% CI)	0.77 [0.61, 0.96]
1.5.7 vorozole 2.5 mg	1	452	Odds Ratio (M-H, Random, 95% CI)	1.25 [0.83, 1.88]
1.6 Objective response (randomised)	31	10422	Odds Ratio (M-H, Random, 95% CI)	0.90 [0.78, 1.03]
1.6.1 aminoglutethimide (any dose)	11	1765	Odds Ratio (M-H, Random, 95% CI)	0.89 [0.66, 1.20]
1.6.2 formestane 250 mg	3	1133	Odds Ratio (M-H, Random, 95% CI)	1.20 [0.91, 1.60]
1.6.3 anastrozole 1 mg	4	2626	Odds Ratio (M-H, Random, 95% CI)	0.95 [0.77, 1.17]
1.6.4 exemestane 25 mg	3	1584	Odds Ratio (M-H, Random, 95% CI)	0.68 [0.37, 1.27]
1.6.5 fadrozole 2 mg	5	1080	Odds Ratio (M-H, Random, 95% CI)	1.21 [0.87, 1.69]
1.6.6 letrozole 2.5 mg	4	1782	Odds Ratio (M-H, Random, 95% CI)	0.65 [0.52, 0.82]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.6.7 vorozole 2.5 mg	1	452	Odds Ratio (M-H, Random, 95% CI)	0.67 [0.33, 1.37]

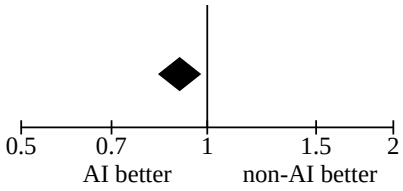
Analysis 1.1. Comparison 1: AI versus non-AI, Outcome 1: Overall survival (reported or calculated)

Study or Subgroup	log[HR]	SE	Weight	HR IV, Fixed, 95% CI
1.1.1 aminoglutethimide (any dose)				
Gale 1994	0.11	0.16	5.8%	1.12 [0.82 , 1.53]
Ingle 1986	-0.2357	0.2502	2.4%	0.79 [0.48 , 1.29]
Rose 1986	-0.0943	0.1491	6.6%	0.91 [0.68 , 1.22]
Russell 1997	-0.1165	0.1708	5.1%	0.89 [0.64 , 1.24]
Subtotal (95% CI)			**19.8%**	**0.94 [0.80 , 1.12]**
Heterogeneity: Chi² = 1.78, df = 3 (P = 0.62); I² = 0%				
Test for overall effect: Z = 0.66 (P = 0.51)				
1.1.2 anastrozole 1 mg				
Bonneterre 2001	-0.0305	0.0931	17.0%	0.97 [0.81 , 1.16]
Buzdar 1996a	-0.2485	0.1277	9.0%	0.78 [0.61 , 1.00]
Milla-Santos 2003	-0.0834	0.1533	6.3%	0.92 [0.68 , 1.24]
Subtotal (95% CI)			**32.4%**	**0.90 [0.79 , 1.03]**
Heterogeneity: Chi² = 1.92, df = 2 (P = 0.38); I² = 0%				
Test for overall effect: Z = 1.51 (P = 0.13)				
1.1.3 exemestane 25 mg				
Kaufmann 2000	-0.1661	0.0805	22.8%	0.85 [0.72 , 0.99]
Subtotal (95% CI)			**22.8%**	**0.85 [0.72 , 0.99]**
Heterogeneity: Not applicable				
Test for overall effect: Z = 2.06 (P = 0.04)				
1.1.4 fadrozole 2 mg				
Bezwoda 1998	0.3001	0.2683	2.0%	1.35 [0.80 , 2.28]
Thuerlimann 1996	-0.0943	0.1887	4.1%	0.91 [0.63 , 1.32]
Subtotal (95% CI)			**6.2%**	**1.04 [0.77 , 1.40]**
Heterogeneity: Chi² = 1.45, df = 1 (P = 0.23); I² = 31%				
Test for overall effect: Z = 0.23 (P = 0.81)				
1.1.5 letrozole 2.5 mg				
Buzdar 2001	-0.0834	0.1203	10.2%	0.92 [0.73 , 1.16]
Dombernowsky 1998	-0.1985	0.1375	7.8%	0.82 [0.63 , 1.07]
Subtotal (95% CI)			**18.0%**	**0.88 [0.73 , 1.05]**
Heterogeneity: Chi² = 0.40, df = 1 (P = 0.53); I² = 0%				
Test for overall effect: Z = 1.47 (P = 0.14)				
1.1.6 vorozole 2.5 mg				
Goss 1999	0.0953	0.4121	0.9%	1.10 [0.49 , 2.47]
Subtotal (95% CI)			**0.9%**	**1.10 [0.49 , 2.47]**
Heterogeneity: Not applicable				
Test for overall effect: Z = 0.23 (P = 0.82)				



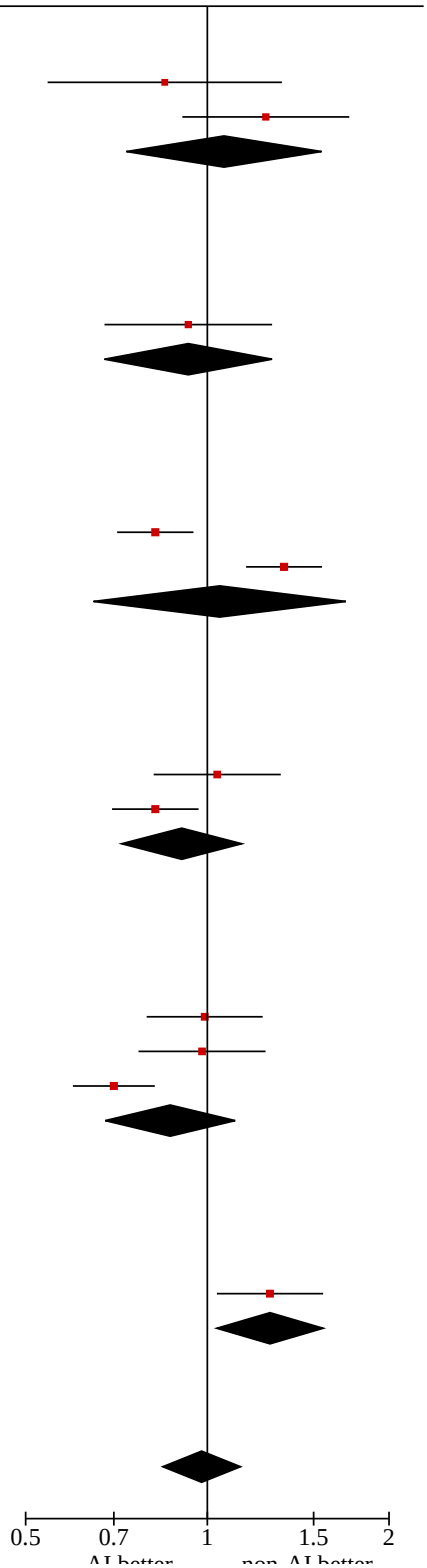
Analysis 1.1. (Continued)

Total (95% CI)	**100.0%**	**0.90 [0.84 , 0.97]**
Heterogeneity: Chi² = 7.60, df = 12 (P = 0.82); I² = 0%		
Test for overall effect: Z = 2.68 (P = 0.007)		
Test for subgroup differences: Chi² = 2.05, df = 5 (P = 0.84), I² = 0%		



Analysis 1.2. Comparison 1: AI versus non-AI, Outcome 2: Progression-free survival (reported or calculated)

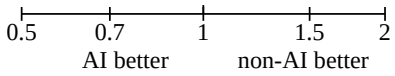
Study or Subgroup	log[HR]	SE	Weight	HR IV, Random, 95% CI
1.2.1 aminoglutethimide (any dose)				
Ingle 1986	-0.1625	0.228	5.7%	0.85 [0.54 , 1.33]
Russell 1997	0.2231	0.1624	7.6%	1.25 [0.91 , 1.72]
Subtotal (95% CI)			**13.3%**	**1.07 [0.73 , 1.55]**
Heterogeneity: Tau² = 0.04; Chi² = 1.90, df = 1 (P = 0.17); I² = 47%				
Test for overall effect: Z = 0.33 (P = 0.74)				
1.2.2 formestane 250 mg				
Thuerlimann 1997	-0.0726	0.1631	7.6%	0.93 [0.68 , 1.28]
Subtotal (95% CI)			**7.6%**	**0.93 [0.68 , 1.28]**
Heterogeneity: Not applicable				
Test for overall effect: Z = 0.45 (P = 0.66)				
1.2.3 anastrozole 1 mg				
Bonneterre 2001	-0.1985	0.0743	10.6%	0.82 [0.71 , 0.95]
Mauriac 2003	0.2927	0.0739	10.6%	1.34 [1.16 , 1.55]
Subtotal (95% CI)			**21.2%**	**1.05 [0.65 , 1.70]**
Heterogeneity: Tau² = 0.12; Chi² = 21.97, df = 1 (P < 0.00001); I² = 95%				
Test for overall effect: Z = 0.19 (P = 0.85)				
1.2.4 exemestane 25 mg				
Chia 2008	0.0377	0.1237	9.0%	1.04 [0.81 , 1.32]
Kaufmann 2000	-0.1985	0.0842	10.3%	0.82 [0.70 , 0.97]
Subtotal (95% CI)			**19.3%**	**0.91 [0.72 , 1.14]**
Heterogeneity: Tau² = 0.02; Chi² = 2.49, df = 1 (P = 0.11); I² = 60%				
Test for overall effect: Z = 0.84 (P = 0.40)				
1.2.5 letrozole 2.5 mg				
Buzdar 2001	-0.0101	0.1129	9.4%	0.99 [0.79 , 1.24]
Dombernowsky 1998	-0.0202	0.1236	9.0%	0.98 [0.77 , 1.25]
Mourisden 2001	-0.3567	0.0797	10.5%	0.70 [0.60 , 0.82]
Subtotal (95% CI)			**28.8%**	**0.87 [0.68 , 1.11]**
Heterogeneity: Tau² = 0.04; Chi² = 8.80, df = 2 (P = 0.01); I² = 77%				
Test for overall effect: Z = 1.12 (P = 0.26)				
1.2.6 vorozole 2.5 mg				
Goss 1999	0.239	0.1034	9.7%	1.27 [1.04 , 1.56]
Subtotal (95% CI)			**9.7%**	**1.27 [1.04 , 1.56]**
Heterogeneity: Not applicable				
Test for overall effect: Z = 2.31 (P = 0.02)				
Total (95% CI)			**100.0%**	**0.98 [0.84 , 1.13]**
Heterogeneity: Tau² = 0.05; Chi² = 54.88, df = 10 (P < 0.00001); I² = 82%				
Test for overall effect: Z = 0.29 (P = 0.77)				
Test for subgroup differences: Chi² = 7.52, df = 5 (P = 0.18), I² = 33.5%				



Analysis 1.2. (Continued)

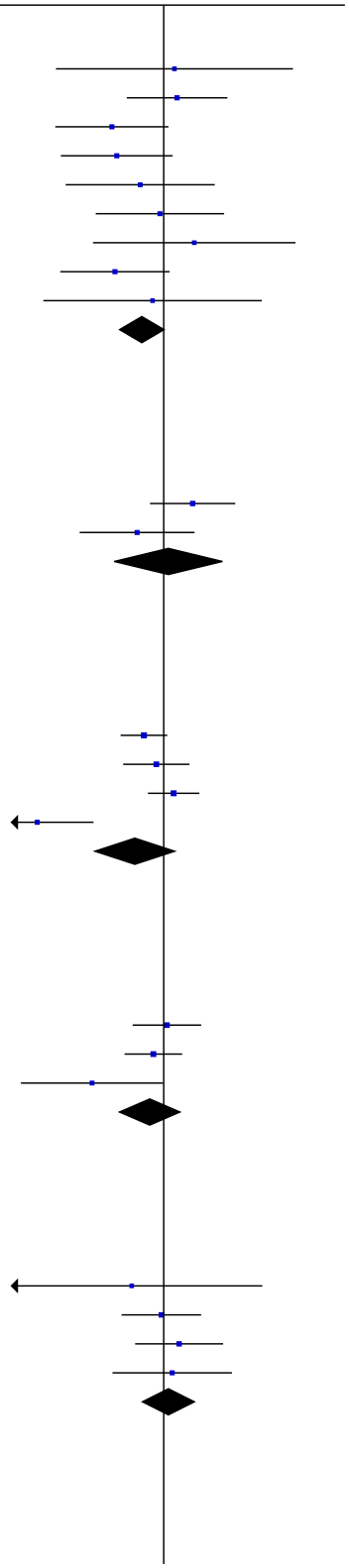
Test for overall effect: Z = 0.29 (P = 0.77)

Test for subgroup differences: Chi² = 7.52, df = 5 (P = 0.18), I² = 33.5%



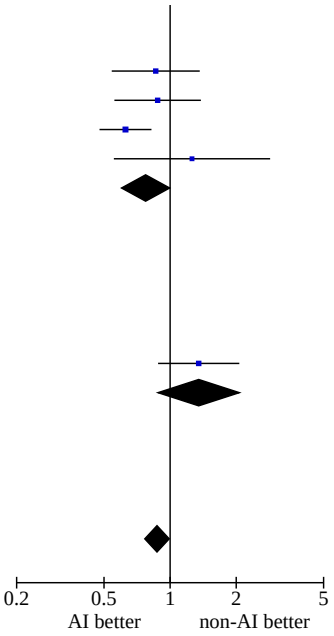
Analysis 1.3. Comparison 1: AI versus non-AI, Outcome 3: Clinical benefit (assessable)

Study or Subgroup	non-AI Events	non-AI Total	AI Events	AI Total	Weight	Odds Ratio M-H, Random, 95% CI
1.3.1 aminoglutethimide (any dose)						
Alonso-Munoz 1988	28	34	25	31	0.9%	1.12 [0.32 , 3.92]
Canney 1988	61	112	54	106	3.6%	1.15 [0.68 , 1.96]
Gale 1994	71	108	83	108	3.1%	0.58 [0.32 , 1.05]
Garcia-Giralt 1992	78	112	98	124	3.2%	0.61 [0.34 , 1.10]
Ingle 1986	21	49	25	51	2.1%	0.78 [0.35 , 1.72]
Lundgren 1989	49	74	51	76	2.6%	0.96 [0.49 , 1.89]
Mercer 1993	13	29	10	27	1.2%	1.38 [0.47 , 4.03]
Powles 1984	55	99	67	99	3.3%	0.60 [0.33 , 1.06]
Samonis 1994	18	27	18	26	1.1%	0.89 [0.28 , 2.82]
Subtotal (95% CI)		**644**		**648**	**21.1%**	**0.79 [0.63 , 1.00]**
Total events:	394		431			
Heterogeneity: Tau² = 0.00; Chi² = 6.33, df = 8 (P = 0.61); I² = 0%						
Test for overall effect: Z = 1.94 (P = 0.05)						
1.3.2 formestane 250 mg						
Perez Carrion 1994	124	175	111	173	4.4%	1.36 [0.87 , 2.13]
Thuerlimann 1997	46	83	56	90	3.0%	0.75 [0.41 , 1.39]
Subtotal (95% CI)		**258**		**263**	**7.5%**	**1.05 [0.59 , 1.86]**
Total events:	170		167			
Heterogeneity: Tau² = 0.10; Chi² = 2.32, df = 1 (P = 0.13); I² = 57%						
Test for overall effect: Z = 0.17 (P = 0.87)						
1.3.3 anastrozole 1 mg						
Bonneterre 2001	265	510	292	511	7.2%	0.81 [0.63 , 1.04]
Buzdar 1996a	102	253	111	263	5.6%	0.93 [0.65 , 1.31]
Mauriac 2003	186	428	173	423	6.8%	1.11 [0.85 , 1.46]
Milla-Santos 2003	65	117	100	121	3.1%	0.26 [0.14 , 0.48]
Subtotal (95% CI)		**1308**		**1318**	**22.7%**	**0.74 [0.48 , 1.12]**
Total events:	618		676			
Heterogeneity: Tau² = 0.15; Chi² = 19.04, df = 3 (P = 0.0003); I² = 84%						
Test for overall effect: Z = 1.42 (P = 0.16)						
1.3.4 exemestane 25 mg						
Chia 2008	87	270	85	270	5.5%	1.03 [0.72 , 1.49]
Kaufmann 2000	135	366	133	337	6.3%	0.90 [0.66 , 1.22]
Paridaens 2003	25	57	35	56	2.2%	0.47 [0.22 , 0.99]
Subtotal (95% CI)		**693**		**663**	**14.0%**	**0.86 [0.63 , 1.19]**
Total events:	247		253			
Heterogeneity: Tau² = 0.03; Chi² = 3.45, df = 2 (P = 0.18); I² = 42%						
Test for overall effect: Z = 0.91 (P = 0.36)						
1.3.5 fadrozole 2 mg						
Bezwoda 1998	4	50	5	46	0.8%	0.71 [0.18 , 2.84]
Buzdar 1996b	65	184	70	195	4.7%	0.98 [0.64 , 1.49]
Buzdar 1996c	61	148	56	150	4.3%	1.18 [0.74 , 1.87]
Thuerlimann 1996	81	106	77	103	2.9%	1.09 [0.58 , 2.06]
Subtotal (95% CI)		**488**		**494**	**12.7%**	**1.05 [0.80 , 1.38]**
Total events:	211		208			
Heterogeneity: Tau² = 0.00; Chi² = 0.67, df = 3 (P = 0.88); I² = 0%						
Test for overall effect: Z = 0.35 (P = 0.72)						
1.3.6 letrozole 2.5 mg						



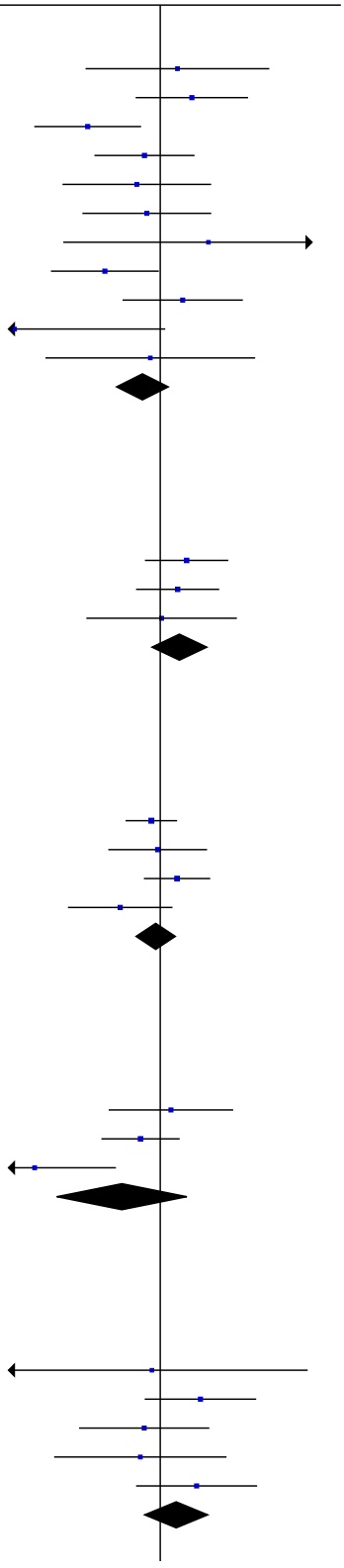
Analysis 1.3. (Continued)

1.3.6 letrozole 2.5 mg						
Buzdar 2001	47	180	53	182	4.3%	0.86 [0.54 , 1.36]
Dombernowsky 1998	60	166	60	153	4.4%	0.88 [0.56 , 1.38]
Mourisden 2001	173	423	221	421	6.8%	0.63 [0.48 , 0.82]
Schmid 2001	19	60	14	52	1.9%	1.26 [0.55 , 2.85]
Subtotal (95% CI)		**829**		**808**	**17.4%**	**0.77 [0.60 , 1.00]**
Total events:	299		348			
Heterogeneity: $Tau^2 = 0.02$; $Chi^2 = 4.00$, df = 3 (P = 0.26); $I^2 = 25\%$						
Test for overall effect: Z = 2.00 (P = 0.05)						
1.3.7 vorozole 2.5 mg						
Goss 1999	71	185	60	190	4.7%	1.35 [0.88 , 2.07]
Subtotal (95% CI)		**185**		**190**	**4.7%**	**1.35 [0.88 , 2.07]**
Total events:	71		60			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.38 (P = 0.17)						
Total (95% CI)		**4405**		**4384**	**100.0%**	**0.87 [0.77 , 0.99]**
Total events:	2010		2143			
Heterogeneity: $Tau^2 = 0.04$; $Chi^2 = 46.41$, df = 26 (P = 0.008); $I^2 = 44\%$						
Test for overall effect: Z = 2.11 (P = 0.03)						
Test for subgroup differences: $Chi^2 = 8.31$, df = 6 (P = 0.22), $I^2 = 27.8\%$						



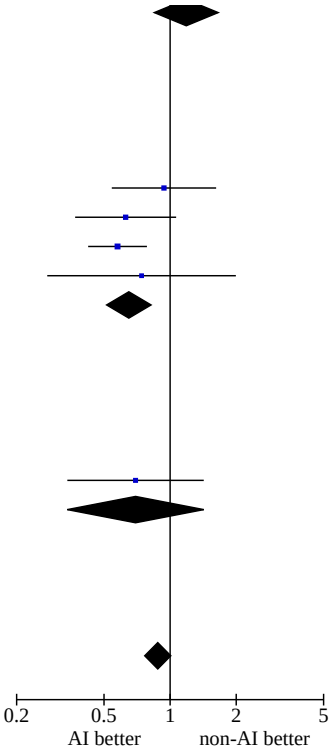
Analysis 1.4. Comparison 1: AI versus non-AI, Outcome 4: Objective response (assessable)

Study or Subgroup	non-AI Events	non-AI Total	AI Events	AI Total	Weight	Odds Ratio M-H, Random, 95% CI
1.4.1 aminoglutethimide (any dose)						
Alonso-Munoz 1988	18	34	15	31	1.6%	1.20 [0.45 , 3.18]
Canney 1988	35	112	26	106	3.4%	1.40 [0.77 , 2.54]
Gale 1994	30	108	49	108	3.6%	0.46 [0.26 , 0.82]
Garcia-Giralt 1992	39	112	48	124	3.9%	0.85 [0.50 , 1.44]
Ingle 1986	21	49	25	51	2.3%	0.78 [0.35 , 1.72]
Lundgren 1989	23	74	26	76	2.8%	0.87 [0.44 , 1.72]
Mercer 1993	5	29	3	27	0.7%	1.67 [0.36 , 7.77]
Powles 1984	34	99	48	99	3.6%	0.56 [0.31 , 0.98]
Rose 1986	32	94	24	83	3.1%	1.27 [0.67 , 2.40]
Russell 1997	2	32	10	42	0.7%	0.21 [0.04 , 1.05]
Samonis 1994	9	27	10	28	1.3%	0.90 [0.30 , 2.74]
Subtotal (95% CI)		**770**		**775**	**27.1%**	**0.83 [0.63 , 1.09]**
Total events:	248		284			
Heterogeneity: Tau² = 0.07; Chi² = 14.78, df = 10 (P = 0.14); I² = 32%						
Test for overall effect: Z = 1.35 (P = 0.18)						
1.4.2 formestane 250 mg						
Freue 2000	55	237	45	242	4.8%	1.32 [0.85 , 2.06]
Perez Carrion 1994	65	175	57	173	4.8%	1.20 [0.77 , 1.87]
Thuerlimann 1997	14	83	15	90	2.2%	1.01 [0.46 , 2.25]
Subtotal (95% CI)		**495**		**505**	**11.9%**	**1.23 [0.92 , 1.64]**
Total events:	134		117			
Heterogeneity: Tau² = 0.00; Chi² = 0.34, df = 2 (P = 0.84); I² = 0%						
Test for overall effect: Z = 1.37 (P = 0.17)						
1.4.3 anastrozole 1 mg						
Bonneterre 2001	138	510	148	511	7.0%	0.91 [0.69 , 1.20]
Buzdar 1996a	31	253	33	263	4.0%	0.97 [0.58 , 1.64]
Mauriac 2003	82	428	70	423	5.9%	1.20 [0.84 , 1.70]
Milla-Santos 2003	31	117	43	121	3.7%	0.65 [0.38 , 1.14]
Subtotal (95% CI)		**1308**		**1318**	**20.6%**	**0.95 [0.77 , 1.17]**
Total events:	282		294			
Heterogeneity: Tau² = 0.01; Chi² = 3.48, df = 3 (P = 0.32); I² = 14%						
Test for overall effect: Z = 0.46 (P = 0.65)						
1.4.4 exemestane 25 mg						
Chia 2008	20	270	18	270	3.0%	1.12 [0.58 , 2.17]
Kaufmann 2000	50	366	55	337	5.1%	0.81 [0.54 , 1.23]
Paridaens 2003	10	57	25	56	2.0%	0.26 [0.11 , 0.62]
Subtotal (95% CI)		**693**		**663**	**10.1%**	**0.67 [0.33 , 1.33]**
Total events:	80		98			
Heterogeneity: Tau² = 0.26; Chi² = 7.20, df = 2 (P = 0.03); I² = 72%						
Test for overall effect: Z = 1.16 (P = 0.25)						
1.4.5 fadrozole 2 mg						
Bezwoda 1998	3	50	3	46	0.6%	0.91 [0.18 , 4.78]
Buzdar 1996b	30	184	22	195	3.4%	1.53 [0.85 , 2.77]
Buzdar 1996c	17	148	20	150	2.8%	0.84 [0.42 , 1.68]
Falkson 1996	17	38	18	36	1.8%	0.81 [0.32 , 2.02]
Thuerlimann 1996	29	106	21	103	3.1%	1.47 [0.77 , 2.79]
Subtotal (95% CI)		**526**		**530**	**11.8%**	**1.18 [0.85 , 1.65]**
Total events:	96		84			



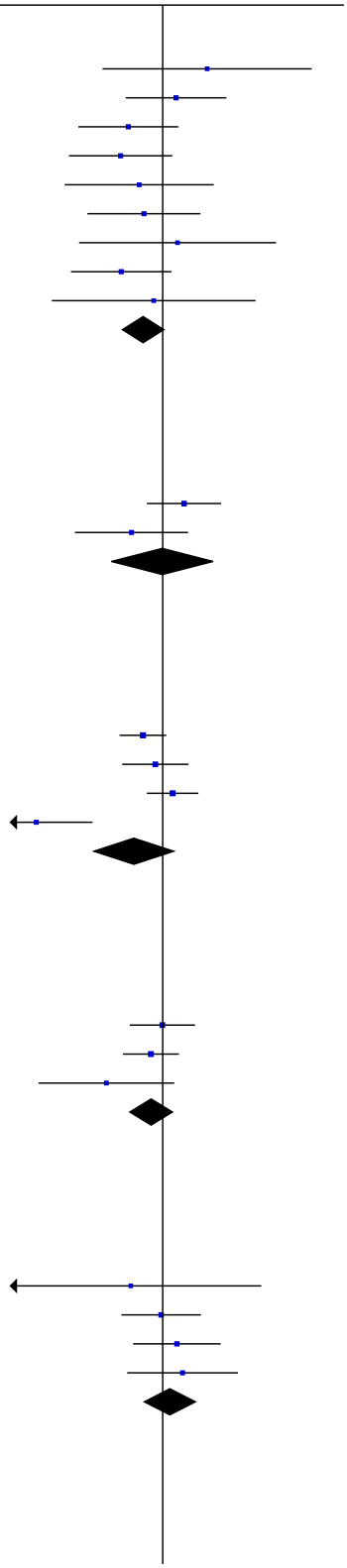
Analysis 1.4. (Continued)

Subtotal (95% CI)		**526**		**530**	**11.8%**	**1.18 [0.85 , 1.65]**
Total events:	96		84			
Heterogeneity: Tau² = 0.00; Chi² = 2.85, df = 4 (P = 0.58); I² = 0%						
Test for overall effect: Z = 0.99 (P = 0.32)						
1.4.6 letrozole 2.5 mg						
Buzdar 2001	30	180	32	182	3.8%	0.94 [0.54 , 1.62]
Dombernowsky 1998	31	166	41	153	3.9%	0.63 [0.37 , 1.07]
Mourisden 2001	92	423	137	421	6.5%	0.58 [0.42 , 0.78]
Schmid 2001	9	60	10	52	1.6%	0.74 [0.28 , 1.99]
Subtotal (95% CI)		**829**		**808**	**15.8%**	**0.65 [0.51 , 0.82]**
Total events:	162		220			
Heterogeneity: Tau² = 0.00; Chi² = 2.39, df = 3 (P = 0.49); I² = 0%						
Test for overall effect: Z = 3.65 (P = 0.0003)						
1.4.7 vorozole 2.5 mg						
Goss 1999	14	185	20	190	2.7%	0.70 [0.34 , 1.42]
Subtotal (95% CI)		**185**		**190**	**2.7%**	**0.70 [0.34 , 1.42]**
Total events:	14		20			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.99 (P = 0.32)						
Total (95% CI)		**4806**		**4789**	**100.0%**	**0.88 [0.77 , 1.01]**
Total events:	1016		1117			
Heterogeneity: Tau² = 0.05; Chi² = 48.15, df = 30 (P = 0.02); I² = 38%						
Test for overall effect: Z = 1.86 (P = 0.06)						
Test for subgroup differences: Chi² = 16.36, df = 6 (P = 0.01), I² = 63.3%						



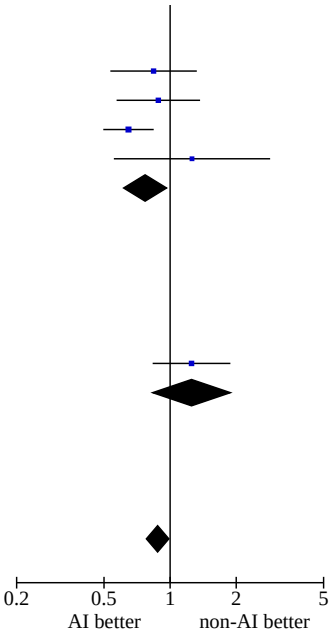
Analysis 1.5. Comparison 1: AI versus non-AI, Outcome 5: Clinical benefit (randomised)

Study or Subgroup	non-AI Events	non-AI Total	AI Events	AI Total	Weight	Odds Ratio M-H, Random, 95% CI
1.5.1 aminoglutethimide (any dose)						
Alonso-Munoz 1988	28	35	25	35	1.0%	1.60 [0.53 , 4.84]
Canney 1988	61	112	54	106	3.4%	1.15 [0.68 , 1.96]
Gale 1994	71	119	83	122	3.4%	0.70 [0.41 , 1.18]
Garcia-Giralt 1992	78	119	98	131	3.2%	0.64 [0.37 , 1.11]
Ingle 1986	21	49	25	51	1.8%	0.78 [0.35 , 1.72]
Lundgren 1989	49	90	51	86	2.9%	0.82 [0.45 , 1.49]
Mercer 1993	13	33	10	28	1.1%	1.17 [0.41 , 3.32]
Powles 1984	55	111	67	111	3.4%	0.64 [0.38 , 1.10]
Samonis 1994	18	29	18	28	1.1%	0.91 [0.31 , 2.67]
Subtotal (95% CI)		**697**		**698**	**21.3%**	**0.81 [0.65 , 1.01]**
Total events:	394		431			
Heterogeneity: Tau² = 0.00; Chi² = 5.40, df = 8 (P = 0.71); I² = 0%						
Test for overall effect: Z = 1.86 (P = 0.06)						
1.5.2 formestane 250 mg						
Perez Carrion 1994	124	206	111	203	4.9%	1.25 [0.85 , 1.86]
Thuerlimann 1997	46	86	56	91	2.9%	0.72 [0.40 , 1.31]
Subtotal (95% CI)		**292**		**294**	**7.8%**	**1.00 [0.58 , 1.70]**
Total events:	170		167			
Heterogeneity: Tau² = 0.09; Chi² = 2.32, df = 1 (P = 0.13); I² = 57%						
Test for overall effect: Z = 0.02 (P = 0.99)						
1.5.3 anastrozole 1 mg						
Bonneterre 2001	265	510	292	511	7.3%	0.81 [0.63 , 1.04]
Buzdar 1996a	102	253	111	263	5.5%	0.93 [0.65 , 1.31]
Mauriac 2003	186	428	173	423	6.9%	1.11 [0.85 , 1.46]
Milla-Santos 2003	65	117	100	121	2.9%	0.26 [0.14 , 0.48]
Subtotal (95% CI)		**1308**		**1318**	**22.6%**	**0.74 [0.48 , 1.12]**
Total events:	618		676			
Heterogeneity: Tau² = 0.15; Chi² = 19.04, df = 3 (P = 0.0003); I² = 84%						
Test for overall effect: Z = 1.42 (P = 0.16)						
1.5.4 exemestane 25 mg						
Chia 2008	87	351	85	342	5.6%	1.00 [0.71 , 1.41]
Kaufmann 2000	135	403	133	366	6.4%	0.88 [0.66 , 1.19]
Paridaens 2003	25	60	35	62	2.2%	0.55 [0.27 , 1.13]
Subtotal (95% CI)		**814**		**770**	**14.2%**	**0.88 [0.71 , 1.11]**
Total events:	247		253			
Heterogeneity: Tau² = 0.00; Chi² = 2.13, df = 2 (P = 0.35); I² = 6%						
Test for overall effect: Z = 1.07 (P = 0.28)						
1.5.5 fadrozole 2 mg						
Bezwoda 1998	4	50	5	46	0.7%	0.71 [0.18 , 2.84]
Buzdar 1996b	65	184	70	196	4.5%	0.98 [0.65 , 1.50]
Buzdar 1996c	61	151	56	152	4.0%	1.16 [0.73 , 1.85]
Thuerlimann 1996	81	110	77	111	2.9%	1.23 [0.69 , 2.21]
Subtotal (95% CI)		**495**		**505**	**12.2%**	**1.08 [0.82 , 1.41]**
Total events:	211		208			
Heterogeneity: Tau² = 0.00; Chi² = 0.83, df = 3 (P = 0.84); I² = 0%						
Test for overall effect: Z = 0.55 (P = 0.58)						
1.5.6 letrozole 2.5 mg						



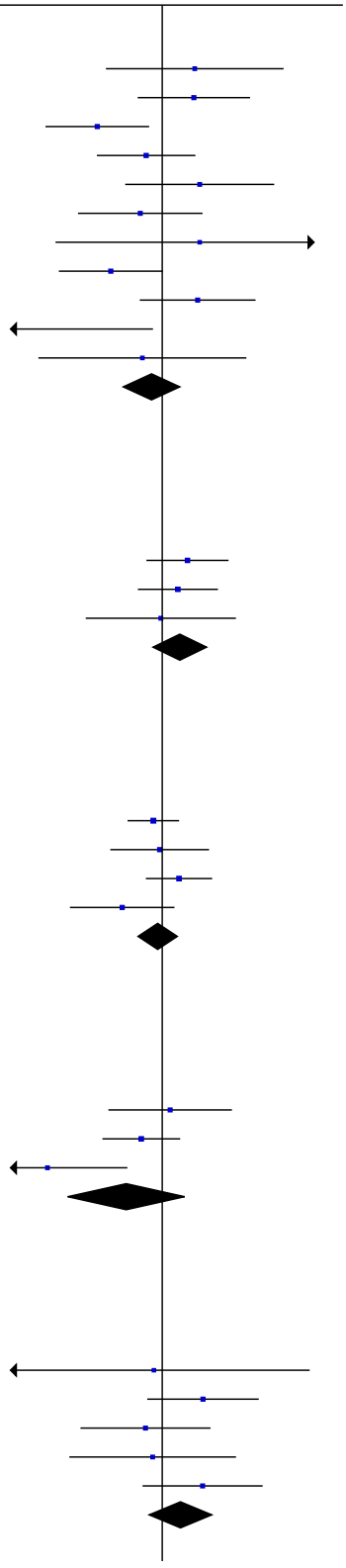
Analysis 1.5. (Continued)

1.5.6 letrozole 2.5 mg						
Buzdar 2001	47	201	53	199	4.2%	0.84 [0.53 , 1.32]
Dombernowsky 1998	60	189	60	174	4.3%	0.88 [0.57 , 1.37]
Mourisden 2001	173	454	221	453	7.0%	0.65 [0.50 , 0.84]
Schmid 2001	19	60	14	52	1.7%	1.26 [0.55 , 2.85]
Subtotal (95% CI)		**904**		**878**	**17.2%**	**0.77 [0.61 , 0.96]**
Total events:	299		348			
Heterogeneity: Tau² = 0.01; Chi² = 3.53, df = 3 (P = 0.32); I² = 15%						
Test for overall effect: Z = 2.29 (P = 0.02)						
1.5.7 vorozole 2.5 mg						
Goss 1999	71	227	60	225	4.7%	1.25 [0.83 , 1.88]
Subtotal (95% CI)		**227**		**225**	**4.7%**	**1.25 [0.83 , 1.88]**
Total events:	71		60			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.08 (P = 0.28)						
Total (95% CI)		**4737**		**4688**	**100.0%**	**0.88 [0.78 , 0.99]**
Total events:	2010		2143			
Heterogeneity: Tau² = 0.03; Chi² = 42.67, df = 26 (P = 0.02); I² = 39%						
Test for overall effect: Z = 2.18 (P = 0.03)						
Test for subgroup differences: Chi² = 7.83, df = 6 (P = 0.25), I² = 23.3%						



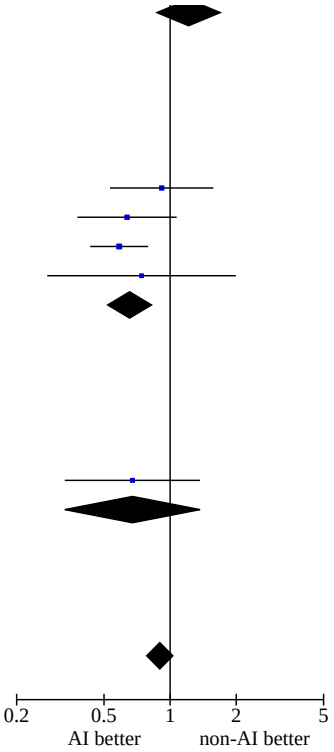
Analysis 1.6. Comparison 1: AI versus non-AI, Outcome 6: Objective response (randomised)

Study or Subgroup	non-AI Events	non-AI Total	AI Events	AI Total	Weight	Odds Ratio M-H, Random, 95% CI
1.6.1 aminoglutethimide (any dose)						
Alonso-Munoz 1988	18	35	15	35	1.7%	1.41 [0.55 , 3.62]
Canney 1988	35	112	26	106	3.4%	1.40 [0.77 , 2.54]
Gale 1994	30	119	49	122	3.7%	0.50 [0.29 , 0.87]
Garcia-Giralt 1992	39	119	48	131	3.9%	0.84 [0.50 , 1.42]
Ingle 1986	25	49	21	51	2.3%	1.49 [0.68 , 3.28]
Lundgren 1989	23	90	26	86	2.9%	0.79 [0.41 , 1.53]
Mercer 1993	5	33	3	28	0.7%	1.49 [0.32 , 6.87]
Powles 1984	34	111	48	111	3.7%	0.58 [0.33 , 1.01]
Rose 1986	32	108	24	107	3.2%	1.46 [0.79 , 2.69]
Russell 1997	2	75	10	80	0.7%	0.19 [0.04 , 0.91]
Samonis 1994	9	29	10	28	1.3%	0.81 [0.27 , 2.44]
Subtotal (95% CI)		**880**		**885**	**27.7%**	**0.89 [0.66 , 1.20]**
Total events:	252		280			
Heterogeneity: Tau² = 0.11; Chi² = 18.08, df = 10 (P = 0.05); I² = 45%						
Test for overall effect: Z = 0.74 (P = 0.46)						
1.6.2 formestane 250 mg						
Freue 2000	55	271	45	276	4.8%	1.31 [0.85 , 2.02]
Perez Carrion 1994	65	206	57	203	4.9%	1.18 [0.77 , 1.80]
Thuerlimann 1997	14	86	15	91	2.2%	0.99 [0.44 , 2.18]
Subtotal (95% CI)		**563**		**570**	**11.9%**	**1.20 [0.91 , 1.60]**
Total events:	134		117			
Heterogeneity: Tau² = 0.00; Chi² = 0.39, df = 2 (P = 0.82); I² = 0%						
Test for overall effect: Z = 1.29 (P = 0.20)						
1.6.3 anastrozole 1 mg						
Bonneterre 2001	138	510	148	511	6.8%	0.91 [0.69 , 1.20]
Buzdar 1996a	31	253	33	263	3.9%	0.97 [0.58 , 1.64]
Mauriac 2003	82	428	70	423	5.8%	1.20 [0.84 , 1.70]
Milla-Santos 2003	31	117	43	121	3.7%	0.65 [0.38 , 1.14]
Subtotal (95% CI)		**1308**		**1318**	**20.1%**	**0.95 [0.77 , 1.17]**
Total events:	282		294			
Heterogeneity: Tau² = 0.01; Chi² = 3.48, df = 3 (P = 0.32); I² = 14%						
Test for overall effect: Z = 0.46 (P = 0.65)						
1.6.4 exemestane 25 mg						
Chia 2008	20	351	18	342	3.0%	1.09 [0.56 , 2.09]
Kaufmann 2000	50	403	55	366	5.0%	0.80 [0.53 , 1.21]
Paridaens 2003	10	60	25	62	2.0%	0.30 [0.13 , 0.69]
Subtotal (95% CI)		**814**		**770**	**10.1%**	**0.68 [0.37 , 1.27]**
Total events:	80		98			
Heterogeneity: Tau² = 0.20; Chi² = 5.96, df = 2 (P = 0.05); I² = 66%						
Test for overall effect: Z = 1.21 (P = 0.23)						
1.6.5 fadrozole 2 mg						
Bezwoda 1998	3	50	3	46	0.6%	0.91 [0.18 , 4.78]
Buzdar 1996b	30	184	22	196	3.4%	1.54 [0.85 , 2.78]
Buzdar 1996c	17	151	20	152	2.8%	0.84 [0.42 , 1.67]
Falkson 1996	17	40	18	40	1.9%	0.90 [0.37 , 2.19]
Thuerlimann 1996	29	110	21	111	3.1%	1.53 [0.81 , 2.90]
Subtotal (95% CI)		**535**		**545**	**11.8%**	**1.21 [0.87 , 1.69]**
Total events:	96		84			



Analysis 1.6. (Continued)

Subtotal (95% CI)		**535**		**545**	**11.8%**	**1.21 [0.87 , 1.69]**
Total events:	96		84			
Heterogeneity: Tau² = 0.00; Chi² = 2.80, df = 4 (P = 0.59); I² = 0%						
Test for overall effect: Z = 1.14 (P = 0.25)						
1.6.6 letrozole 2.5 mg						
Buzdar 2001	30	201	32	199	3.8%	0.92 [0.53 , 1.57]
Dombernowsky 1998	31	189	41	174	4.0%	0.64 [0.38 , 1.07]
Mourisden 2001	92	454	137	453	6.4%	0.59 [0.43 , 0.79]
Schmid 2001	9	60	10	52	1.6%	0.74 [0.28 , 1.99]
Subtotal (95% CI)		**904**		**878**	**15.7%**	**0.65 [0.52 , 0.82]**
Total events:	162		220			
Heterogeneity: Tau² = 0.00; Chi² = 2.05, df = 3 (P = 0.56); I² = 0%						
Test for overall effect: Z = 3.62 (P = 0.0003)						
1.6.7 vorozole 2.5 mg						
Goss 1999	14	227	20	225	2.7%	0.67 [0.33 , 1.37]
Subtotal (95% CI)		**227**		**225**	**2.7%**	**0.67 [0.33 , 1.37]**
Total events:	14		20			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.09 (P = 0.28)						
Total (95% CI)		**5231**		**5191**	**100.0%**	**0.90 [0.78 , 1.03]**
Total events:	1020		1113			
Heterogeneity: Tau² = 0.05; Chi² = 49.68, df = 30 (P = 0.01); I² = 40%						
Test for overall effect: Z = 1.55 (P = 0.12)						
Test for subgroup differences: Chi² = 16.30, df = 6 (P = 0.01), I² = 63.2%						



Comparison 2. AI versus non-AI: Toxicity

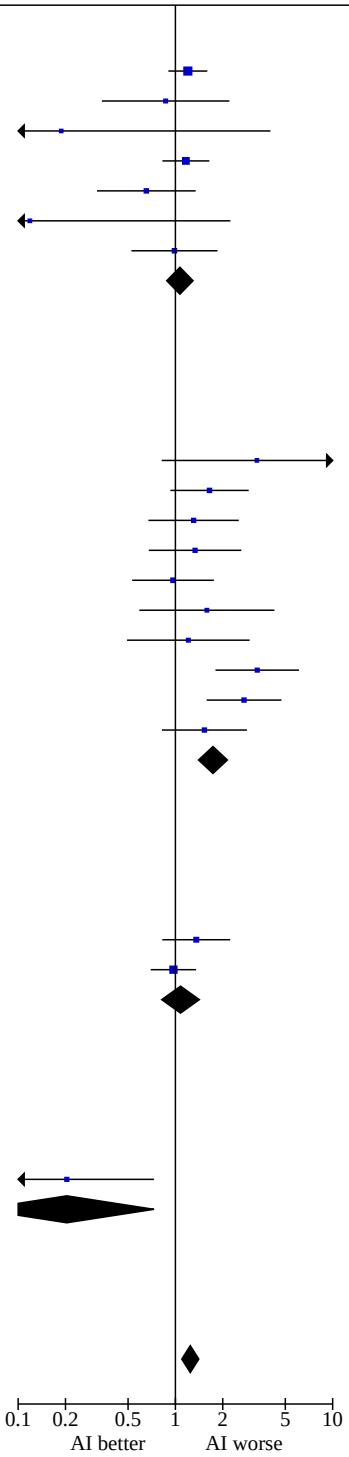
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.1 hot flushes	20	8306	Odds Ratio (M-H, Fixed, 95% CI)	1.24 [1.10, 1.41]
2.1.1 AI versus tamoxifen	7	2616	Odds Ratio (M-H, Fixed, 95% CI)	1.07 [0.88, 1.29]
2.1.2 AI versus megestrol acetate	10	3926	Odds Ratio (M-H, Fixed, 95% CI)	1.73 [1.40, 2.14]
2.1.3 AI versus fulvestrant	2	1546	Odds Ratio (M-H, Fixed, 95% CI)	1.08 [0.82, 1.42]
2.1.4 AI versus medroxyprogesterone acetate	1	218	Odds Ratio (M-H, Fixed, 95% CI)	0.20 [0.06, 0.73]
2.2 nausea	18		Odds Ratio (M-H, Random, 95% CI)	Subtotals only
2.2.1 AI versus tamoxifen	6	2548	Odds Ratio (M-H, Random, 95% CI)	1.29 [0.78, 2.13]
2.2.2 AI versus megestrol acetate	9	3755	Odds Ratio (M-H, Random, 95% CI)	1.77 [1.33, 2.35]
2.2.3 AI versus medroxyprogesterone acetate	1	53	Odds Ratio (M-H, Random, 95% CI)	8.19 [0.40, 166.83]
2.2.4 AI versus fulvestrant	2	1539	Odds Ratio (M-H, Random, 95% CI)	1.01 [0.77, 1.32]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.3 vomiting	8		Odds Ratio (M-H, Fixed, 95% CI)	Subtotals only
2.3.1 AI versus tamoxifen	2	1239	Odds Ratio (M-H, Fixed, 95% CI)	1.23 [0.79, 1.90]
2.3.2 AI versus megestrol acetate	5	2319	Odds Ratio (M-H, Fixed, 95% CI)	2.03 [1.42, 2.90]
2.3.3 AI versus fulvestrant	1	846	Odds Ratio (M-H, Fixed, 95% CI)	0.90 [0.60, 1.35]
2.4 diarrhoea	10		Odds Ratio (M-H, Fixed, 95% CI)	Subtotals only
2.4.1 AI versus tamoxifen	3	2149	Odds Ratio (M-H, Fixed, 95% CI)	1.64 [1.06, 2.55]
2.4.2 AI versus megestrol acetate	5	1961	Odds Ratio (M-H, Fixed, 95% CI)	1.48 [1.02, 2.13]
2.4.3 AI versus fulvestrant	2	1090	Odds Ratio (M-H, Fixed, 95% CI)	1.23 [0.79, 1.90]
2.5 rash	15		Odds Ratio (M-H, Random, 95% CI)	Subtotals only
2.5.1 AI versus tamoxifen	4	711	Odds Ratio (M-H, Random, 95% CI)	33.61 [4.71, 239.97]
2.5.2 AI versus megestrol acetate	8	3219	Odds Ratio (M-H, Random, 95% CI)	2.06 [0.92, 4.62]
2.5.3 AI versus medroxyprogesterone acetate	2	271	Odds Ratio (M-H, Random, 95% CI)	36.80 [3.35, 404.73]
2.5.4 AI versus fulvestrant	1	397	Odds Ratio (M-H, Random, 95% CI)	1.39 [0.77, 2.50]
2.6 vaginal bleeding	6		Odds Ratio (M-H, Fixed, 95% CI)	Subtotals only
2.6.1 AI versus tamoxifen	1	1017	Odds Ratio (M-H, Fixed, 95% CI)	0.45 [0.16, 1.32]
2.6.2 AI versus megestrol acetate	3	1462	Odds Ratio (M-H, Fixed, 95% CI)	0.22 [0.10, 0.45]
2.6.3 AI versus medroxyprogesterone acetate	2	271	Odds Ratio (M-H, Fixed, 95% CI)	0.13 [0.02, 0.71]
2.7 thromboembolic	6		Odds Ratio (M-H, Fixed, 95% CI)	Subtotals only
2.7.1 AI versus tamoxifen	2	1228	Odds Ratio (M-H, Fixed, 95% CI)	0.48 [0.27, 0.85]
2.7.2 AI versus megestrol acetate	3	863	Odds Ratio (M-H, Fixed, 95% CI)	0.54 [0.26, 1.10]
2.7.3 AI versus fulvestrant	1	846	Odds Ratio (M-H, Fixed, 95% CI)	1.14 [0.56, 2.31]
2.8 arthralgia	6		Odds Ratio (M-H, Fixed, 95% CI)	Subtotals only
2.8.1 AI versus tamoxifen	2	1031	Odds Ratio (M-H, Fixed, 95% CI)	1.14 [0.81, 1.60]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.8.2 AI versus megestrol acetate	4	1439	Odds Ratio (M-H, Fixed, 95% CI)	1.40 [0.98, 2.00]

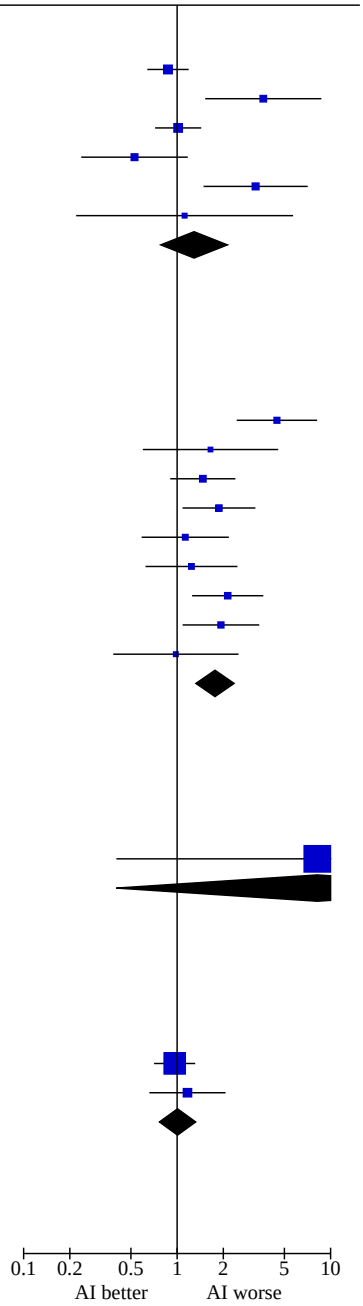
Analysis 2.1. Comparison 2: AI versus non-AI: Toxicity, Outcome 1: hot flushes

Study or Subgroup	AI Events	AI Total	comparison Events	comparison Total	Weight	Odds Ratio M-H, Fixed, 95% CI
2.1.1 AI versus tamoxifen						
Bonneterre 2001	134	506	118	511	19.4%	1.20 [0.90 , 1.60]
Falkson 1996	13	37	15	39	2.1%	0.87 [0.34 , 2.20]
Ingle 1986	46	48	49	49	0.6%	0.19 [0.01 , 4.02]
Mourisden 2001	84	455	74	455	13.6%	1.17 [0.83 , 1.64]
Paridaens 2003	24	62	29	59	4.1%	0.65 [0.32 , 1.35]
Rose 1986	0	87	4	97	1.0%	0.12 [0.01 , 2.24]
Thuerlimann 1996	25	104	26	107	4.4%	0.99 [0.52 , 1.85]
Subtotal (95% CI)		**1299**		**1317**	**45.2%**	**1.07 [0.88 , 1.29]**
Total events:	326		315			
Heterogeneity: Chi² = 6.31, df = 6 (P = 0.39); I² = 5%						
Test for overall effect: Z = 0.67 (P = 0.50)						
2.1.2 AI versus megestrol acetate						
Bezwoda 1998	8	46	3	50	0.5%	3.30 [0.82 , 13.30]
Buzdar 1996a	34	262	21	253	4.2%	1.65 [0.93 , 2.92]
Buzdar 1996b	23	196	17	184	3.5%	1.31 [0.67 , 2.53]
Buzdar 1996c	22	152	17	151	3.3%	1.33 [0.68 , 2.63]
Buzdar 2001	24	199	25	201	4.9%	0.97 [0.53 , 1.76]
Dombernowsky 1998	10	174	7	189	1.4%	1.59 [0.59 , 4.26]
Freue 2000	11	276	9	271	2.0%	1.21 [0.49 , 2.96]
Goss 1999	44	195	16	198	2.8%	3.31 [1.80 , 6.11]
Kaufmann 2000	45	358	20	400	3.7%	2.73 [1.58 , 4.72]
Thuerlimann 1997	39	90	27	81	3.6%	1.53 [0.82 , 2.85]
Subtotal (95% CI)		**1948**		**1978**	**29.9%**	**1.73 [1.40 , 2.14]**
Total events:	260		162			
Heterogeneity: Chi² = 13.58, df = 9 (P = 0.14); I² = 34%						
Test for overall effect: Z = 5.09 (P < 0.00001)						
2.1.3 AI versus fulvestrant						
Chia 2008	39	342	31	358	6.0%	1.36 [0.83 , 2.23]
Mauriac 2003	87	423	89	423	15.9%	0.97 [0.70 , 1.35]
Subtotal (95% CI)		**765**		**781**	**22.0%**	**1.08 [0.82 , 1.42]**
Total events:	126		120			
Heterogeneity: Chi² = 1.20, df = 1 (P = 0.27); I² = 17%						
Test for overall effect: Z = 0.53 (P = 0.59)						
2.1.4 AI versus medroxyprogesterone acetate						
Canney 1988	3	106	14	112	3.0%	0.20 [0.06 , 0.73]
Subtotal (95% CI)		**106**		**112**	**3.0%**	**0.20 [0.06 , 0.73]**
Total events:	3		14			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.44 (P = 0.01)						
Total (95% CI)		**4118**		**4188**	**100.0%**	**1.24 [1.10 , 1.41]**
Total events:	715		611			
Heterogeneity: Chi² = 40.22, df = 19 (P = 0.003); I² = 53%						
Test for overall effect: Z = 3.42 (P = 0.0006)						
Test for subgroup differences: Chi² = 20.59, df = 3 (P = 0.0001), I² = 85.4%						



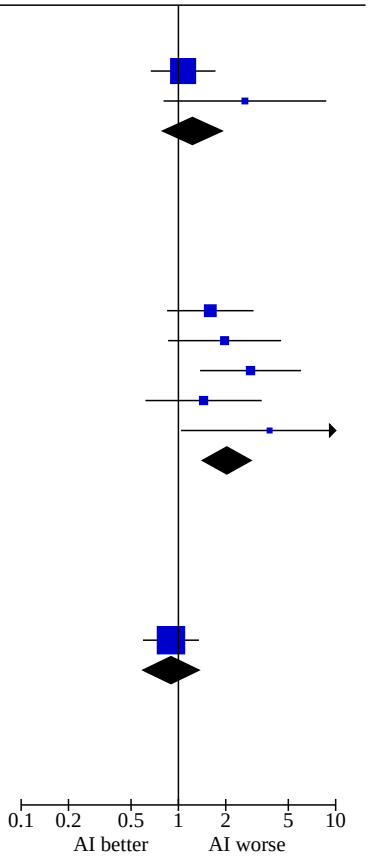
Analysis 2.2. Comparison 2: AI versus non-AI: Toxicity, Outcome 2: nausea

Study or Subgroup	AI Events	AI Total	comparison Events	comparison Total	Weight	Odds Ratio M-H, Random, 95% CI
2.2.1 AI versus tamoxifen						
Bonneterre 2001	94	506	106	511	23.7%	0.87 [0.64 , 1.19]
Ingle 1986	27	48	12	46	14.6%	3.64 [1.52 , 8.70]
Mourisden 2001	78	455	77	455	23.2%	1.02 [0.72 , 1.44]
Paridaens 2003	14	62	21	59	15.7%	0.53 [0.24 , 1.17]
Powles 1984	27	111	10	111	15.9%	3.25 [1.49 , 7.09]
Rose 1986	3	87	3	97	6.9%	1.12 [0.22 , 5.70]
Subtotal (95% CI)		**1269**		**1279**	**100.0%**	**1.29 [0.78 , 2.13]**
Total events:	243		229			
Heterogeneity: Tau² = 0.25; Chi² = 20.15, df = 5 (P = 0.001); I² = 75%						
Test for overall effect: Z = 1.00 (P = 0.32)						
2.2.2 AI versus megestrol acetate						
Buzdar 1996c	55	152	17	151	12.0%	4.47 [2.44 , 8.17]
Bezwoda 1998	11	46	8	50	6.1%	1.65 [0.60 , 4.55]
Buzdar 1996a	46	262	32	253	14.6%	1.47 [0.90 , 2.40]
Buzdar 1996b	43	196	24	184	13.2%	1.87 [1.08 , 3.24]
Buzdar 2001	21	199	19	201	10.9%	1.13 [0.59 , 2.17]
Dombernowsky 1998	19	174	17	189	10.3%	1.24 [0.62 , 2.47]
Goss 1999	46	195	25	198	13.5%	2.14 [1.25 , 3.64]
Kaufmann 2000	33	358	20	400	12.6%	1.93 [1.09 , 3.43]
Freue 2000	9	276	9	271	6.8%	0.98 [0.38 , 2.51]
Subtotal (95% CI)		**1858**		**1897**	**100.0%**	**1.77 [1.33 , 2.35]**
Total events:	283		171			
Heterogeneity: Tau² = 0.08; Chi² = 14.57, df = 8 (P = 0.07); I² = 45%						
Test for overall effect: Z = 3.89 (P = 0.0001)						
2.2.3 AI versus medroxyprogesterone acetate						
Samonis 1994	3	26	0	27	100.0%	8.19 [0.40 , 166.83]
Subtotal (95% CI)		**26**		**27**	**100.0%**	**8.19 [0.40 , 166.83]**
Total events:	3		0			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.37 (P = 0.17)						
2.2.4 AI versus fulvestrant						
Mauriac 2003	107	423	110	423	77.4%	0.96 [0.71 , 1.31]
Chia 2008	27	342	24	351	22.6%	1.17 [0.66 , 2.07]
Subtotal (95% CI)		**765**		**774**	**100.0%**	**1.01 [0.77 , 1.32]**
Total events:	134		134			
Heterogeneity: Tau² = 0.00; Chi² = 0.34, df = 1 (P = 0.56); I² = 0%						
Test for overall effect: Z = 0.05 (P = 0.96)						



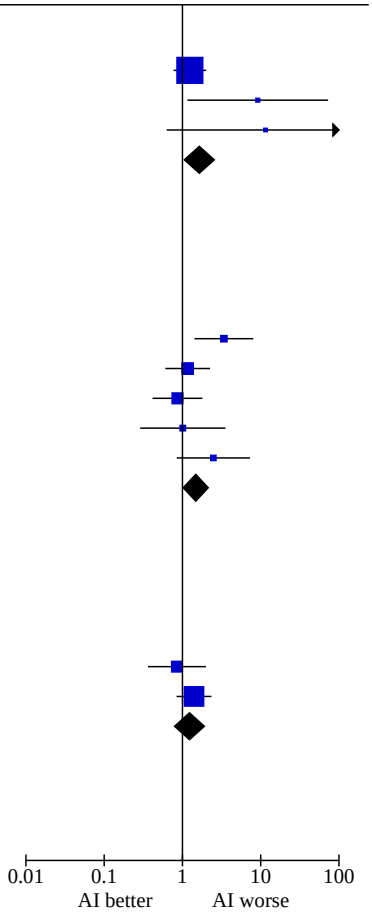
Analysis 2.3. Comparison 2: AI versus non-AI: Toxicity, Outcome 3: vomiting

Study or Subgroup	AI Events	AI Total	non-AI Events	non-AI Total	Weight	Odds Ratio M-H, Fixed, 95% CI
2.3.1 AI versus tamoxifen						
Bonneterre 2001	38	506	36	511	90.1%	1.07 [0.67 , 1.72]
Powles 1984	10	111	4	111	9.9%	2.65 [0.80 , 8.71]
Subtotal (95% CI)		**617**		**622**	**100.0%**	**1.23 [0.79 , 1.90]**
Total events:	48		40			
Heterogeneity: Chi² = 1.92, df = 1 (P = 0.17); I² = 48%						
Test for overall effect: Z = 0.92 (P = 0.36)						
2.3.2 AI versus megestrol acetate						
Buzdar 1996a	27	262	17	253	34.8%	1.59 [0.85 , 3.00]
Buzdar 1996b	18	196	9	184	18.9%	1.97 [0.86 , 4.50]
Buzdar 1996c	28	152	11	151	20.2%	2.87 [1.37 , 6.01]
Dombernowsky 1998	13	174	10	189	19.9%	1.45 [0.62 , 3.39]
Kaufmann 2000	10	358	3	400	6.2%	3.80 [1.04 , 13.93]
Subtotal (95% CI)		**1142**		**1177**	**100.0%**	**2.03 [1.42 , 2.90]**
Total events:	96		50			
Heterogeneity: Chi² = 2.92, df = 4 (P = 0.57); I² = 0%						
Test for overall effect: Z = 3.90 (P < 0.0001)						
2.3.3 AI versus fulvestrant						
Mauriac 2003	50	423	55	423	100.0%	0.90 [0.60 , 1.35]
Subtotal (95% CI)		**423**		**423**	**100.0%**	**0.90 [0.60 , 1.35]**
Total events:	50		55			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.52 (P = 0.60)						



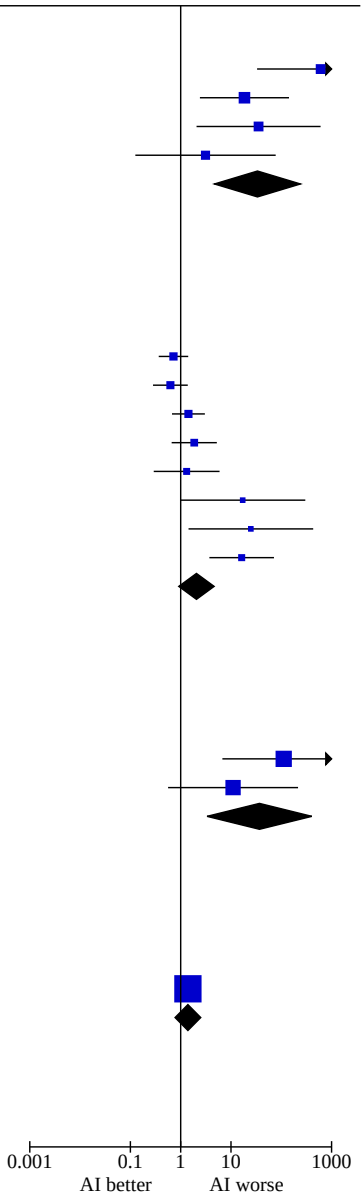
Analysis 2.4. Comparison 2: AI versus non-AI: Toxicity, Outcome 4: diarrhoea

Study or Subgroup	AI Events	AI Total	comparison Events	comparison Total	Weight	Odds Ratio M-H, Fixed, 95% CI
2.4.1 AI versus tamoxifen						
Bonneterre 2001	40	506	33	511	95.4%	1.24 [0.77 , 2.01]
Mourisden 2001	9	455	1	455	3.1%	9.16 [1.16 , 72.61]
Powles 1984	5	111	0	111	1.5%	11.52 [0.63 , 210.82]
Subtotal (95% CI)		**1072**		**1077**	**100.0%**	**1.64 [1.06 , 2.55]**
Total events:	54		34			
Heterogeneity: Chi² = 5.67, df = 2 (P = 0.06); I² = 65%						
Test for overall effect: Z = 2.21 (P = 0.03)						
2.4.2 AI versus megestrol acetate						
Buzdar 1996a	23	262	7	253	13.6%	3.38 [1.42 , 8.03]
Buzdar 1996b	22	196	18	184	34.6%	1.17 [0.60 , 2.25]
Buzdar 1996c	15	152	17	151	32.2%	0.86 [0.41 , 1.80]
Buzdar 2001	5	199	5	201	10.2%	1.01 [0.29 , 3.55]
Dombernowsky 1998	11	174	5	189	9.4%	2.48 [0.85 , 7.30]
Subtotal (95% CI)		**983**		**978**	**100.0%**	**1.48 [1.02 , 2.13]**
Total events:	76		52			
Heterogeneity: Chi² = 7.33, df = 4 (P = 0.12); I² = 45%						
Test for overall effect: Z = 2.09 (P = 0.04)						
2.4.3 AI versus fulvestrant						
Chia 2008	10	342	12	351	31.8%	0.85 [0.36 , 2.00]
Mauriac 2003	40	193	32	204	68.2%	1.41 [0.84 , 2.35]
Subtotal (95% CI)		**535**		**555**	**100.0%**	**1.23 [0.79 , 1.90]**
Total events:	50		44			
Heterogeneity: Chi² = 0.98, df = 1 (P = 0.32); I² = 0%						
Test for overall effect: Z = 0.92 (P = 0.36)						



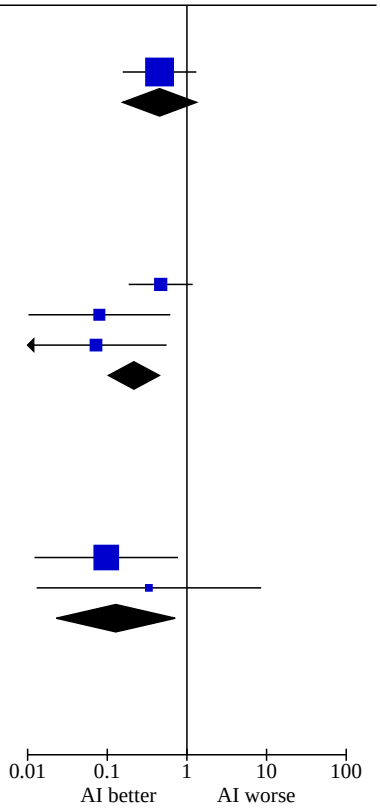
Analysis 2.5. Comparison 2: AI versus non-AI: Toxicity, Outcome 5: rash

Study or Subgroup	AI Events	AI Total	comparison Events	comparison Total	Weight	Odds Ratio M-H, Random, 95% CI
2.5.1 AI versus tamoxifen						
Ingle 1986	48	48	6	46	23.4%	604.38 [33.04 , 11056.23]
Powles 1984	16	111	1	111	31.6%	18.53 [2.41 , 142.31]
Rose 1986	13	87	0	97	24.0%	35.34 [2.07 , 604.03]
Thuerlimann 1996	1	104	0	107	21.0%	3.12 [0.13 , 77.36]
Subtotal (95% CI)		**350**		**361**	**100.0%**	**33.61 [4.71 , 239.97]**
Total events:	78		7			
Heterogeneity: Tau² = 2.10; Chi² = 6.33, df = 3 (P = 0.10); I² = 53%						
Test for overall effect: Z = 3.50 (P = 0.0005)						
2.5.2 AI versus megestrol acetate						
Buzdar 1996a	16	262	21	253	17.2%	0.72 [0.37 , 1.41]
Buzdar 1996b	11	196	16	184	16.4%	0.62 [0.28 , 1.38]
Buzdar 1996c	18	152	13	151	16.7%	1.43 [0.67 , 3.02]
Dombernowsky 1998	10	174	6	189	14.8%	1.86 [0.66 , 5.23]
Freue 2000	4	276	3	271	11.6%	1.31 [0.29 , 5.93]
Kaufmann 2000	7	358	0	400	5.7%	17.09 [0.97 , 300.32]
Lundgren 1989	10	86	0	90	5.7%	24.84 [1.43 , 430.91]
Russell 1997	24	88	2	89	11.8%	16.31 [3.72 , 71.53]
Subtotal (95% CI)		**1592**		**1627**	**100.0%**	**2.06 [0.92 , 4.62]**
Total events:	100		61			
Heterogeneity: Tau² = 0.87; Chi² = 27.30, df = 7 (P = 0.0003); I² = 74%						
Test for overall effect: Z = 1.75 (P = 0.08)						
2.5.3 AI versus medroxyprogesterone acetate						
Canney 1988	35	106	0	112	52.1%	111.71 [6.75 , 1849.91]
Samonis 1994	4	26	0	27	47.9%	11.00 [0.56 , 215.35]
Subtotal (95% CI)		**132**		**139**	**100.0%**	**36.80 [3.35 , 404.73]**
Total events:	39		0			
Heterogeneity: Tau² = 0.82; Chi² = 1.38, df = 1 (P = 0.24); I² = 27%						
Test for overall effect: Z = 2.95 (P = 0.003)						
2.5.4 AI versus fulvestrant						
Mauriac 2003	29	193	23	204	100.0%	1.39 [0.77 , 2.50]
Subtotal (95% CI)		**193**		**204**	**100.0%**	**1.39 [0.77 , 2.50]**
Total events:	29		23			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.10 (P = 0.27)						

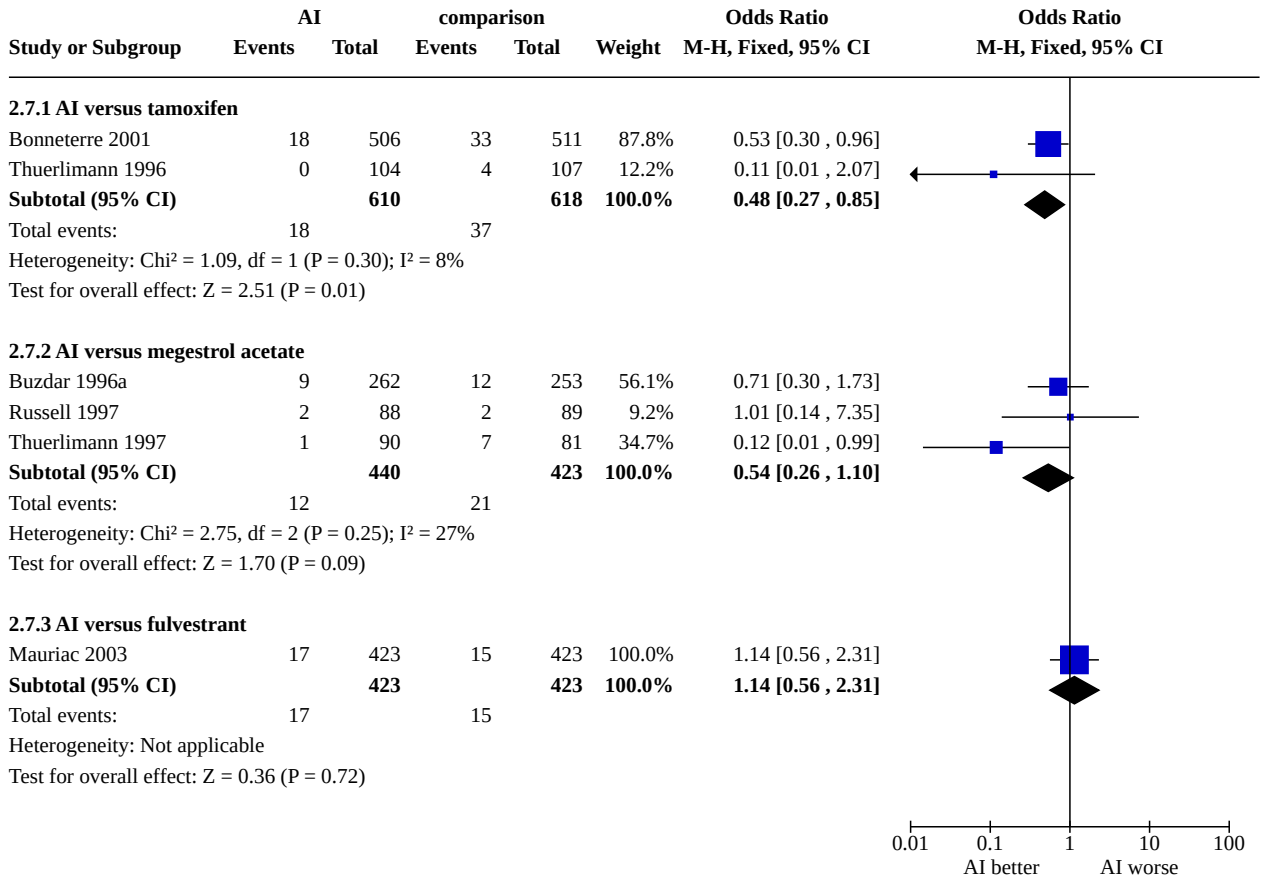


Analysis 2.6. Comparison 2: AI versus non-AI: Toxicity, Outcome 6: vaginal bleeding

Study or Subgroup	AI Events	AI Total	comparison Events	comparison Total	Weight	Odds Ratio M-H, Fixed, 95% CI
2.6.1 AI versus tamoxifen						
Bonneterre 2001	5	506	11	511	100.0%	0.45 [0.16 , 1.32]
Subtotal (95% CI)		**506**		**511**	**100.0%**	**0.45 [0.16 , 1.32]**
Total events:	5		11			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.46 (P = 0.15)						
2.6.2 AI versus megestrol acetate						
Buzdar 1996a	7	262	14	253	35.7%	0.47 [0.19 , 1.18]
Buzdar 2001	1	199	12	201	30.6%	0.08 [0.01 , 0.62]
Freue 2000	1	276	13	271	33.7%	0.07 [0.01 , 0.56]
Subtotal (95% CI)		**737**		**725**	**100.0%**	**0.22 [0.10 , 0.45]**
Total events:	9		39			
Heterogeneity: Chi² = 4.72, df = 2 (P = 0.09); I² = 58%						
Test for overall effect: Z = 4.09 (P < 0.0001)						
2.6.3 AI versus medroxyprogesterone acetate						
Canney 1988	1	106	10	112	87.0%	0.10 [0.01 , 0.77]
Samonis 1994	0	26	1	27	13.0%	0.33 [0.01 , 8.56]
Subtotal (95% CI)		**132**		**139**	**100.0%**	**0.13 [0.02 , 0.71]**
Total events:	1		11			
Heterogeneity: Chi² = 0.40, df = 1 (P = 0.53); I² = 0%						
Test for overall effect: Z = 2.35 (P = 0.02)						

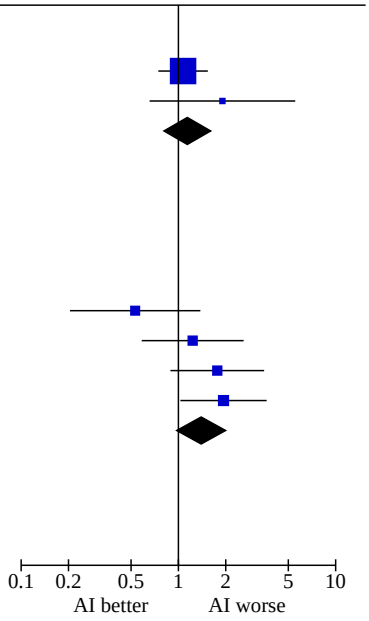


Analysis 2.7. Comparison 2: AI versus non-AI: Toxicity, Outcome 7: thromboembolic



Analysis 2.8. Comparison 2: AI versus non-AI: Toxicity, Outcome 8: arthralgia

Study or Subgroup	AI Events	AI Total	comparison Events	comparison Total	Weight	Odds Ratio M-H, Fixed, 95% CI
2.8.1 AI versus tamoxifen						
Mourisden 2001	71	455	67	455	91.8%	1.07 [0.75 , 1.54]
Paridaens 2003	11	62	6	59	8.2%	1.91 [0.66 , 5.53]
Subtotal (95% CI)		**517**		**514**	**100.0%**	**1.14 [0.81 , 1.60]**
Total events:	82		73			
Heterogeneity: Chi² = 1.01, df = 1 (P = 0.32); I² = 1%						
Test for overall effect: Z = 0.75 (P = 0.46)						
2.8.2 AI versus megestrol acetate						
Buzdar 1996b	7	196	12	184	23.3%	0.53 [0.20 , 1.38]
Buzdar 1996c	17	152	14	151	24.4%	1.23 [0.58 , 2.60]
Dombernowsky 1998	23	174	15	189	24.4%	1.77 [0.89 , 3.51]
Goss 1999	30	195	17	198	27.9%	1.94 [1.03 , 3.64]
Subtotal (95% CI)		**717**		**722**	**100.0%**	**1.40 [0.98 , 2.00]**
Total events:	77		58			
Heterogeneity: Chi² = 5.53, df = 3 (P = 0.14); I² = 46%						
Test for overall effect: Z = 1.83 (P = 0.07)						

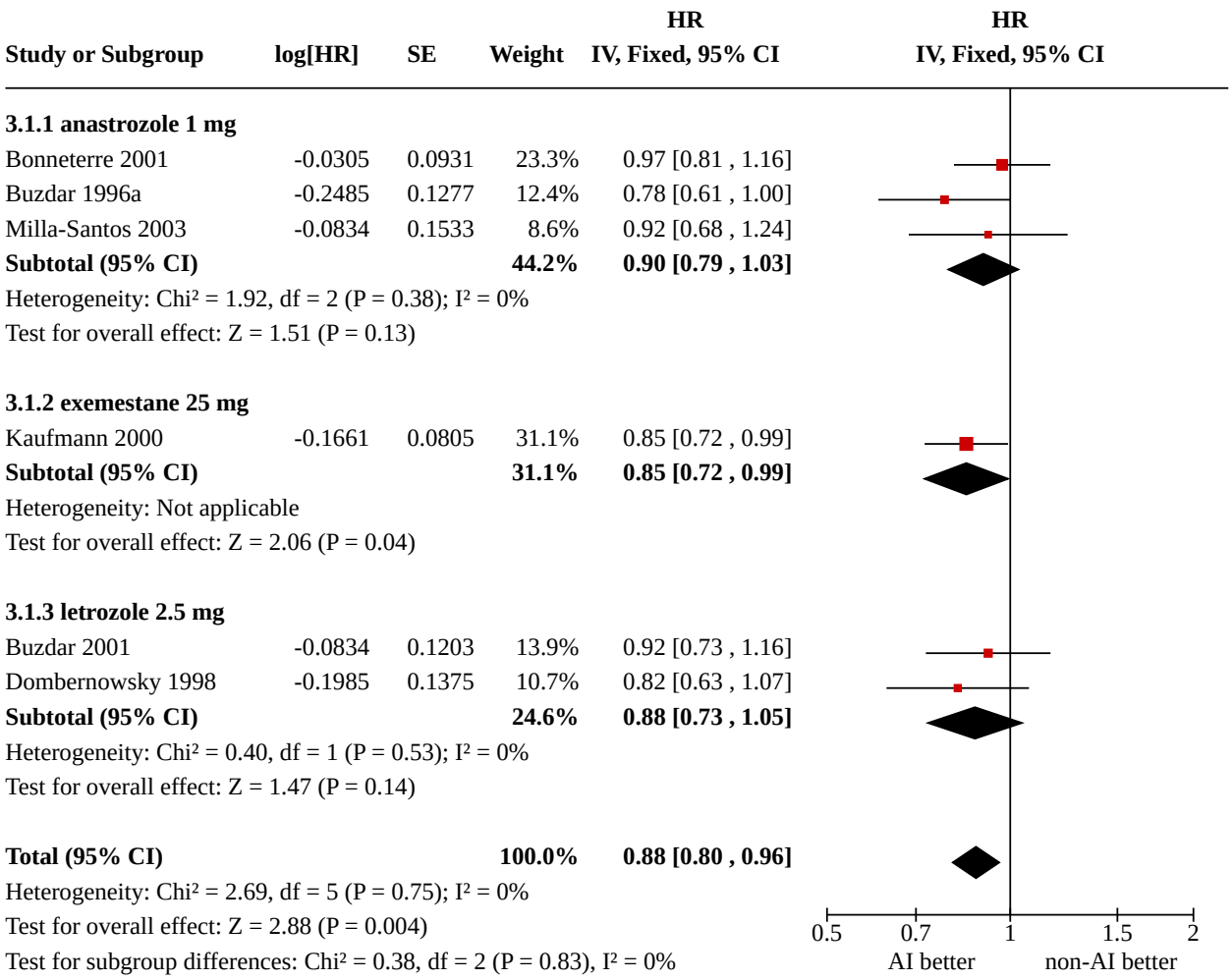


Comparison 3. Current AIs versus non-AI

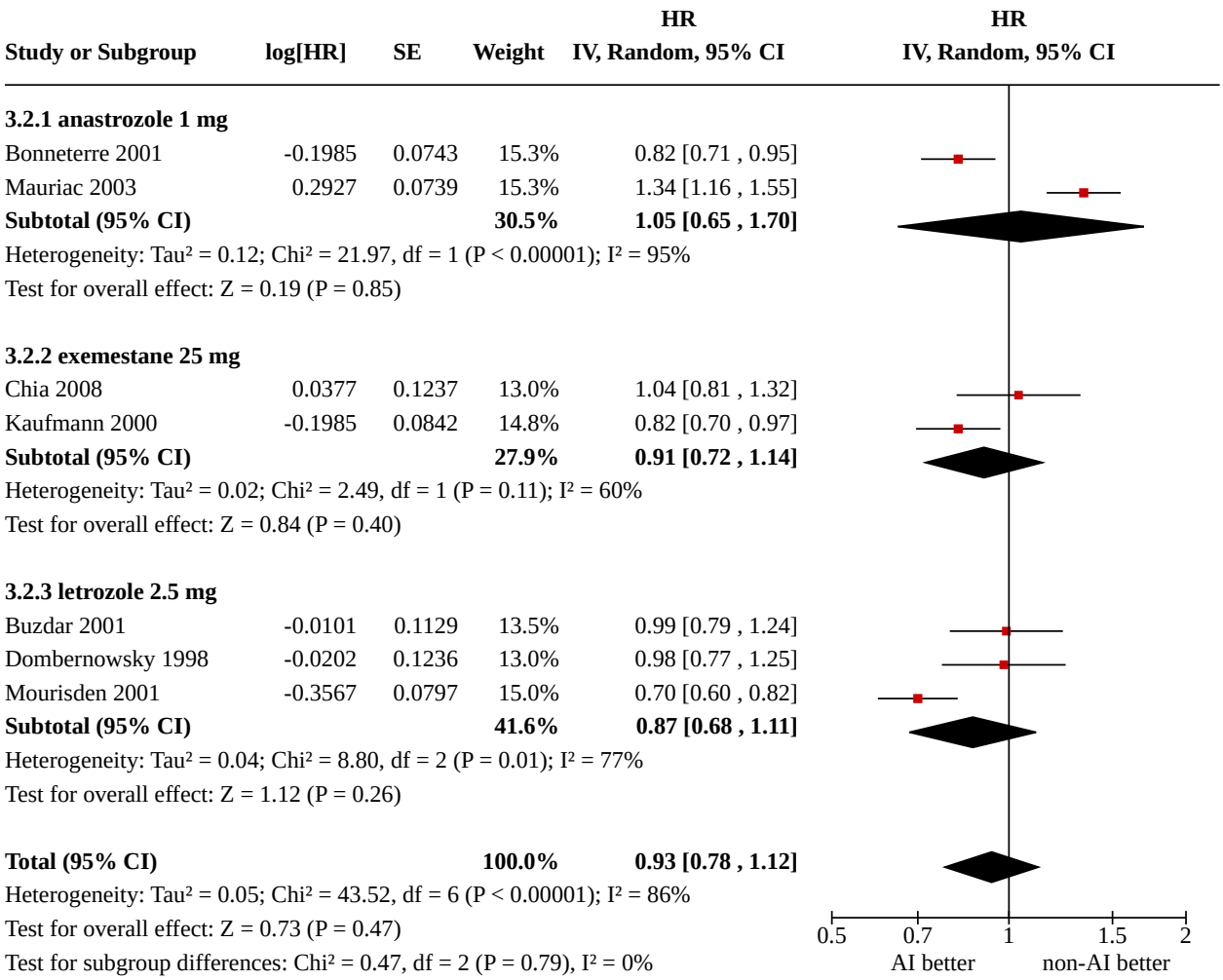
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.1 Overall survival (reported or calculated)	6		HR (IV, Fixed, 95% CI)	0.88 [0.80, 0.96]
3.1.1 anastrozole 1 mg	3		HR (IV, Fixed, 95% CI)	0.90 [0.79, 1.03]
3.1.2 exemestane 25 mg	1		HR (IV, Fixed, 95% CI)	0.85 [0.72, 0.99]
3.1.3 letrozole 2.5 mg	2		HR (IV, Fixed, 95% CI)	0.88 [0.73, 1.05]
3.2 Progression-free survival (reported or calculated)	7		HR (IV, Random, 95% CI)	0.93 [0.78, 1.12]
3.2.1 anastrozole 1 mg	2		HR (IV, Random, 95% CI)	1.05 [0.65, 1.70]
3.2.2 exemestane 25 mg	2		HR (IV, Random, 95% CI)	0.91 [0.72, 1.14]
3.2.3 letrozole 2.5 mg	3		HR (IV, Random, 95% CI)	0.87 [0.68, 1.11]
3.3 Clinical benefit (assessable)	11	5619	Odds Ratio (M-H, Random, 95% CI)	0.80 [0.66, 0.97]
3.3.1 anastrozole 1 mg	4	2626	Odds Ratio (M-H, Random, 95% CI)	0.74 [0.48, 1.12]
3.3.2 exemestane 25 mg	3	1356	Odds Ratio (M-H, Random, 95% CI)	0.86 [0.63, 1.19]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.3.3 letrozole 2.5 mg	4	1637	Odds Ratio (M-H, Random, 95% CI)	0.77 [0.60, 1.00]
3.4 Objective response (assessable)	11	5619	Odds Ratio (M-H, Random, 95% CI)	0.79 [0.65, 0.97]
3.4.1 anastrozole 1 mg	4	2626	Odds Ratio (M-H, Random, 95% CI)	0.95 [0.77, 1.17]
3.4.2 exemestane 25 mg	3	1356	Odds Ratio (M-H, Random, 95% CI)	0.67 [0.33, 1.33]
3.4.3 letrozole 2.5 mg	4	1637	Odds Ratio (M-H, Random, 95% CI)	0.65 [0.51, 0.82]
3.5 Clinical benefit (randomised)	11	5992	Odds Ratio (M-H, Random, 95% CI)	0.81 [0.67, 0.97]
3.5.1 anastrozole 1 mg	4	2626	Odds Ratio (M-H, Random, 95% CI)	0.74 [0.48, 1.12]
3.5.2 exemestane 25 mg	3	1584	Odds Ratio (M-H, Random, 95% CI)	0.88 [0.71, 1.11]
3.5.3 letrozole 2.5 mg	4	1782	Odds Ratio (M-H, Random, 95% CI)	0.77 [0.61, 0.96]
3.6 Objective response (randomised)	11	5992	Odds Ratio (M-H, Random, 95% CI)	0.80 [0.66, 0.96]
3.6.1 anastrozole 1 mg	4	2626	Odds Ratio (M-H, Random, 95% CI)	0.95 [0.77, 1.17]
3.6.2 exemestane 25 mg	3	1584	Odds Ratio (M-H, Random, 95% CI)	0.68 [0.37, 1.27]
3.6.3 letrozole 2.5 mg	4	1782	Odds Ratio (M-H, Random, 95% CI)	0.65 [0.52, 0.82]

Analysis 3.1. Comparison 3: Current AIs versus non-AI, Outcome 1: Overall survival (reported or calculated)

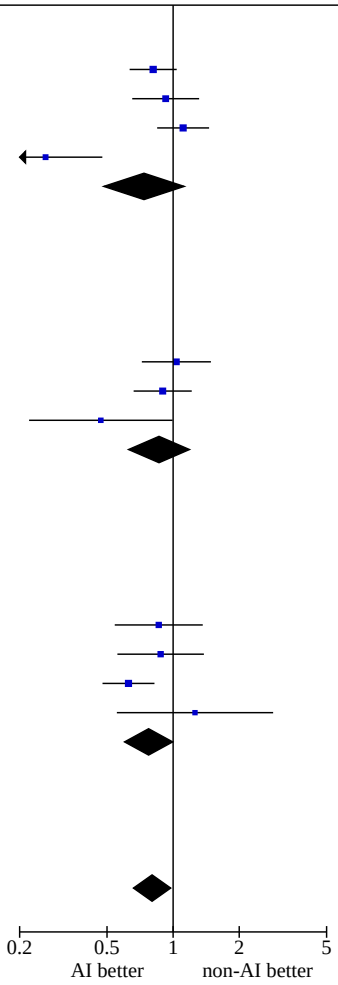


Analysis 3.2. Comparison 3: Current AIs versus non-AI, Outcome 2: Progression-free survival (reported or calculated)

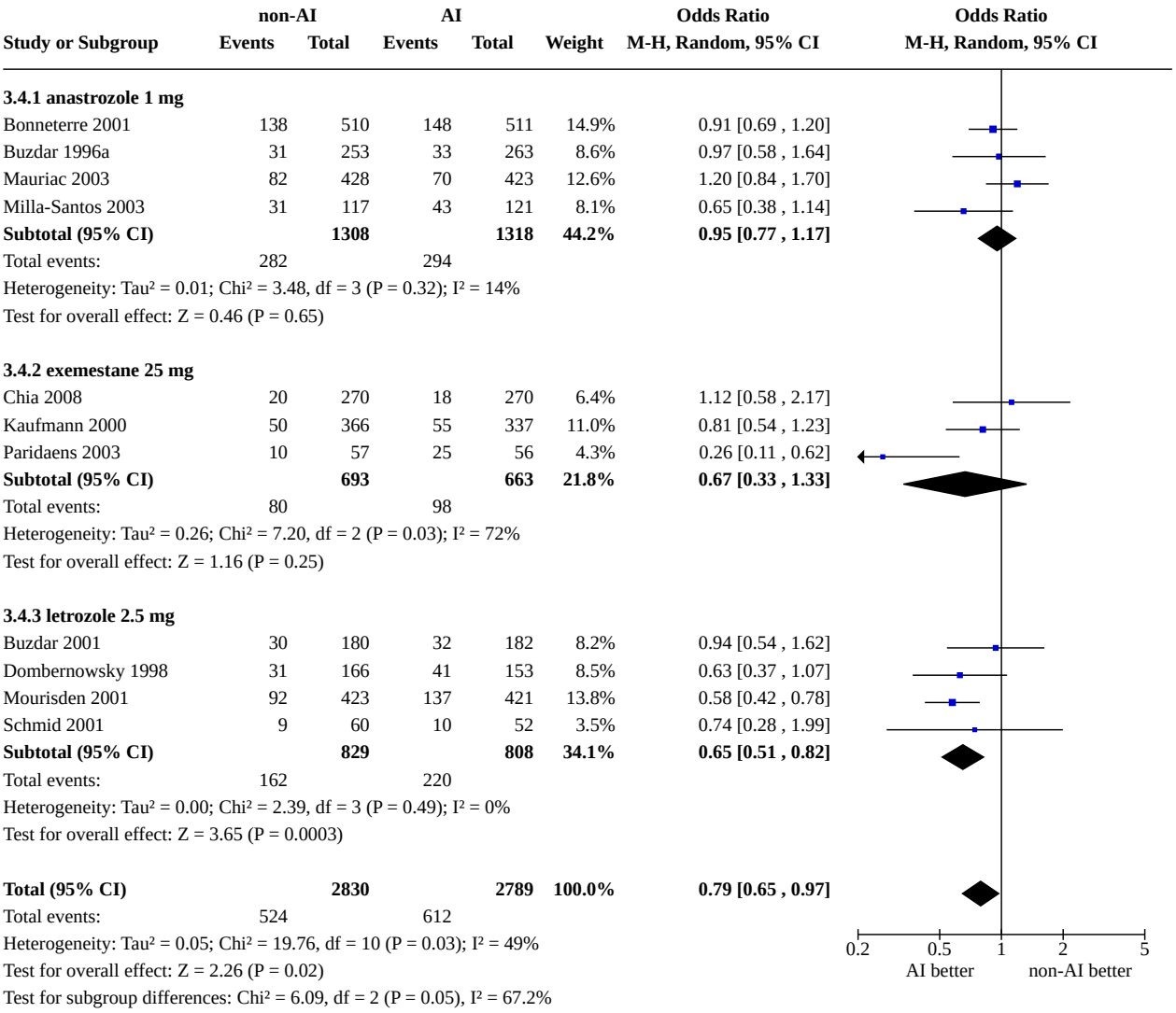


Analysis 3.3. Comparison 3: Current AIs versus non-AI, Outcome 3: Clinical benefit (assessable)

Study or Subgroup	non-AI Events	non-AI Total	AI Events	AI Total	Weight	Odds Ratio M-H, Random, 95% CI
3.3.1 anastrozole 1 mg						
Bonneterre 2001	265	510	292	511	12.5%	0.81 [0.63 , 1.04]
Buzdar 1996a	102	253	111	263	10.3%	0.93 [0.65 , 1.31]
Mauriac 2003	186	428	173	423	11.9%	1.11 [0.85 , 1.46]
Milla-Santos 2003	65	117	100	121	6.3%	0.26 [0.14 , 0.48]
Subtotal (95% CI)		**1308**		**1318**	**41.1%**	**0.74 [0.48 , 1.12]**
Total events:	618		676			
Heterogeneity: Tau² = 0.15; Chi² = 19.04, df = 3 (P = 0.0003); I² = 84%						
Test for overall effect: Z = 1.42 (P = 0.16)						
3.3.2 exemestane 25 mg						
Chia 2008	87	270	85	270	10.1%	1.03 [0.72 , 1.49]
Kaufmann 2000	135	366	133	337	11.3%	0.90 [0.66 , 1.22]
Paridaens 2003	25	57	35	56	4.7%	0.47 [0.22 , 0.99]
Subtotal (95% CI)		**693**		**663**	**26.1%**	**0.86 [0.63 , 1.19]**
Total events:	247		253			
Heterogeneity: Tau² = 0.03; Chi² = 3.45, df = 2 (P = 0.18); I² = 42%						
Test for overall effect: Z = 0.91 (P = 0.36)						
3.3.3 letrozole 2.5 mg						
Buzdar 2001	47	180	53	182	8.3%	0.86 [0.54 , 1.36]
Dombernowsky 1998	60	166	60	153	8.4%	0.88 [0.56 , 1.38]
Mourisden 2001	173	423	221	421	11.9%	0.63 [0.48 , 0.82]
Schmid 2001	19	60	14	52	4.1%	1.26 [0.55 , 2.85]
Subtotal (95% CI)		**829**		**808**	**32.8%**	**0.77 [0.60 , 1.00]**
Total events:	299		348			
Heterogeneity: Tau² = 0.02; Chi² = 4.00, df = 3 (P = 0.26); I² = 25%						
Test for overall effect: Z = 2.00 (P = 0.05)						
Total (95% CI)		**2830**		**2789**	**100.0%**	**0.80 [0.66 , 0.97]**
Total events:	1164		1277			
Heterogeneity: Tau² = 0.06; Chi² = 28.23, df = 10 (P = 0.002); I² = 65%						
Test for overall effect: Z = 2.22 (P = 0.03)						
Test for subgroup differences: Chi² = 0.42, df = 2 (P = 0.81), I² = 0%						

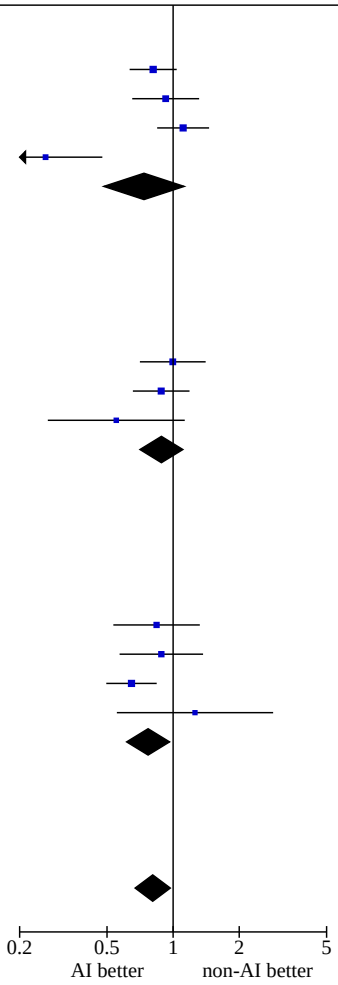


Analysis 3.4. Comparison 3: Current AIs versus non-AI, Outcome 4: Objective response (assessable)

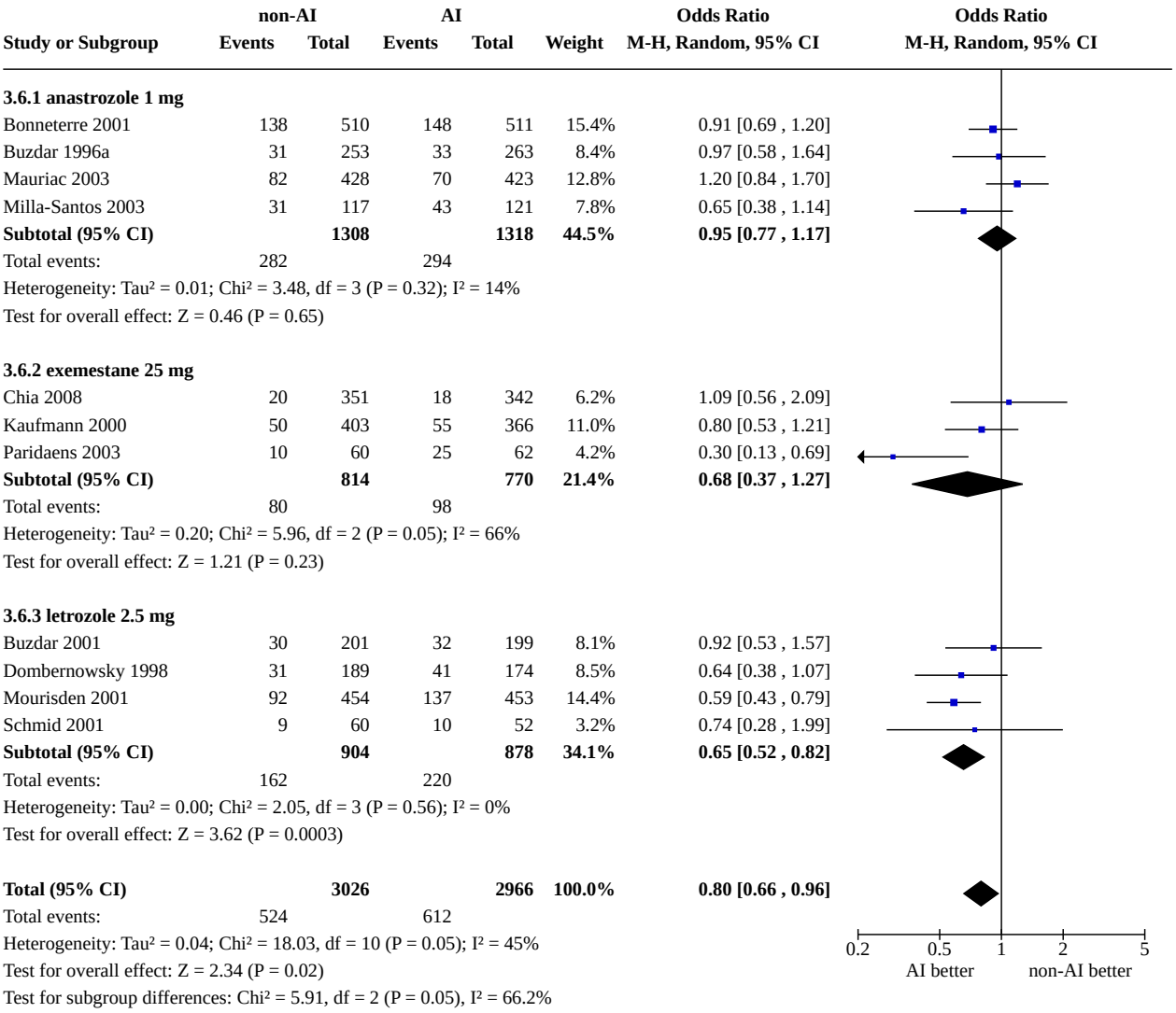


Analysis 3.5. Comparison 3: Current AIs versus non-AI, Outcome 5: Clinical benefit (randomised)

Study or Subgroup	non-AI Events	non-AI Total	AI Events	AI Total	Weight	Odds Ratio M-H, Random, 95% CI
3.5.1 anastrozole 1 mg						
Bonneterre 2001	265	510	292	511	12.6%	0.81 [0.63 , 1.04]
Buzdar 1996a	102	253	111	263	10.2%	0.93 [0.65 , 1.31]
Mauriac 2003	186	428	173	423	12.0%	1.11 [0.85 , 1.46]
Milla-Santos 2003	65	117	100	121	6.0%	0.26 [0.14 , 0.48]
Subtotal (95% CI)		**1308**		**1318**	**40.8%**	**0.74 [0.48 , 1.12]**
Total events:	618		676			
Heterogeneity: Tau² = 0.15; Chi² = 19.04, df = 3 (P = 0.0003); I² = 84%						
Test for overall effect: Z = 1.42 (P = 0.16)						
3.5.2 exemestane 25 mg						
Chia 2008	87	351	85	342	10.4%	1.00 [0.71 , 1.41]
Kaufmann 2000	135	403	133	366	11.4%	0.88 [0.66 , 1.19]
Paridaens 2003	25	60	35	62	4.7%	0.55 [0.27 , 1.13]
Subtotal (95% CI)		**814**		**770**	**26.5%**	**0.88 [0.71 , 1.11]**
Total events:	247		253			
Heterogeneity: Tau² = 0.00; Chi² = 2.13, df = 2 (P = 0.35); I² = 6%						
Test for overall effect: Z = 1.07 (P = 0.28)						
3.5.3 letrozole 2.5 mg						
Buzdar 2001	47	201	53	199	8.2%	0.84 [0.53 , 1.32]
Dombernowsky 1998	60	189	60	174	8.5%	0.88 [0.57 , 1.37]
Mourisden 2001	173	454	221	453	12.2%	0.65 [0.50 , 0.84]
Schmid 2001	19	60	14	52	3.9%	1.26 [0.55 , 2.85]
Subtotal (95% CI)		**904**		**878**	**32.7%**	**0.77 [0.61 , 0.96]**
Total events:	299		348			
Heterogeneity: Tau² = 0.01; Chi² = 3.53, df = 3 (P = 0.32); I² = 15%						
Test for overall effect: Z = 2.29 (P = 0.02)						
Total (95% CI)		**3026**		**2966**	**100.0%**	**0.81 [0.67 , 0.97]**
Total events:	1164		1277			
Heterogeneity: Tau² = 0.05; Chi² = 26.18, df = 10 (P = 0.004); I² = 62%						
Test for overall effect: Z = 2.27 (P = 0.02)						
Test for subgroup differences: Chi² = 0.98, df = 2 (P = 0.61), I² = 0%						



Analysis 3.6. Comparison 3: Current AIs versus non-AI, Outcome 6: Objective response (randomised)

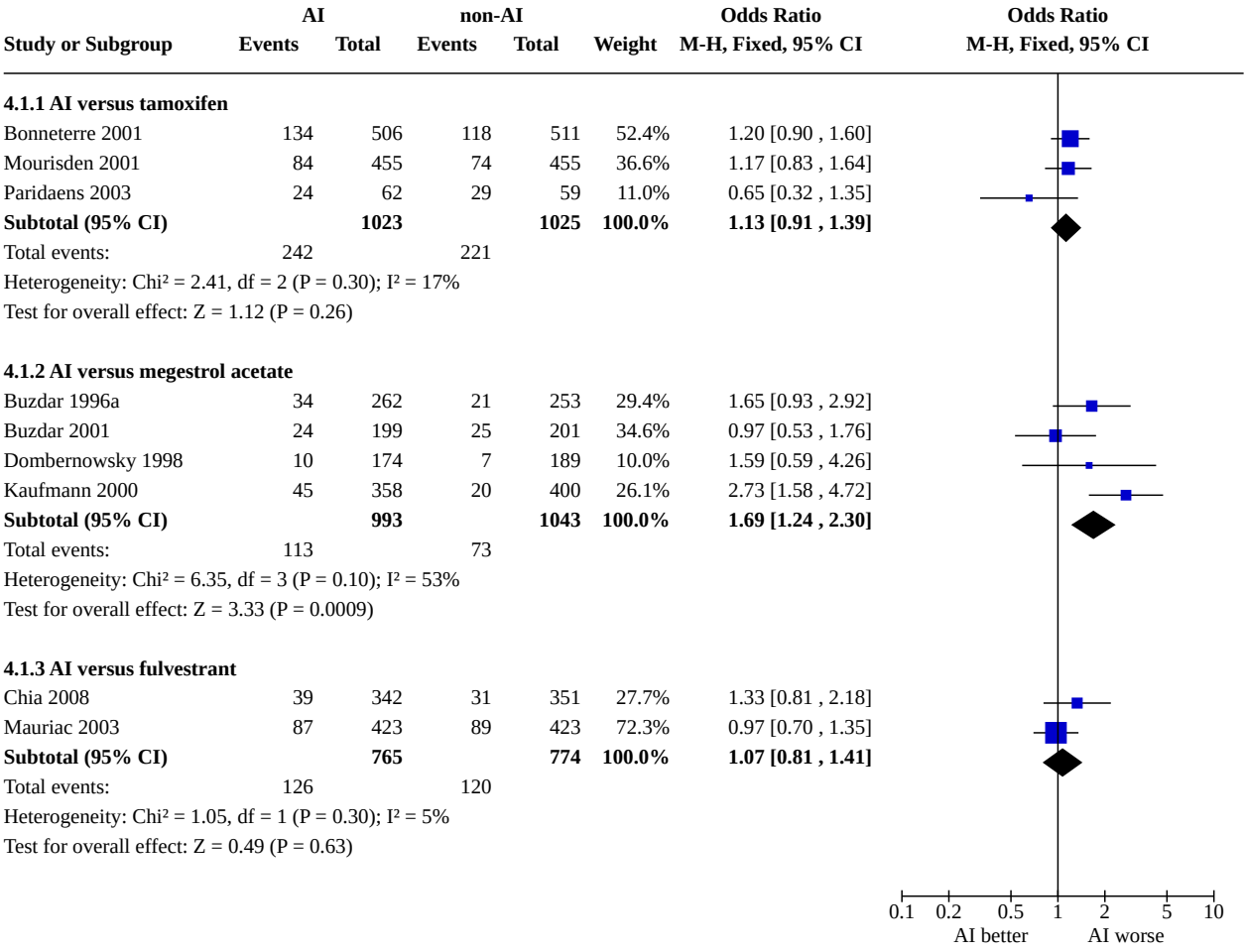


Comparison 4. Current AIs versus non-AI: Toxicity

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
4.1 hot flushes	9		Odds Ratio (M-H, Fixed, 95% CI)	Subtotals only
4.1.1 AI versus tamoxifen	3	2048	Odds Ratio (M-H, Fixed, 95% CI)	1.13 [0.91, 1.39]
4.1.2 AI versus megestrol acetate	4	2036	Odds Ratio (M-H, Fixed, 95% CI)	1.69 [1.24, 2.30]
4.1.3 AI versus fulvestrant	2	1539	Odds Ratio (M-H, Fixed, 95% CI)	1.07 [0.81, 1.41]
4.2 nausea	9		Odds Ratio (M-H, Fixed, 95% CI)	Subtotals only
4.2.1 AI versus tamoxifen	3	2048	Odds Ratio (M-H, Fixed, 95% CI)	0.89 [0.72, 1.11]

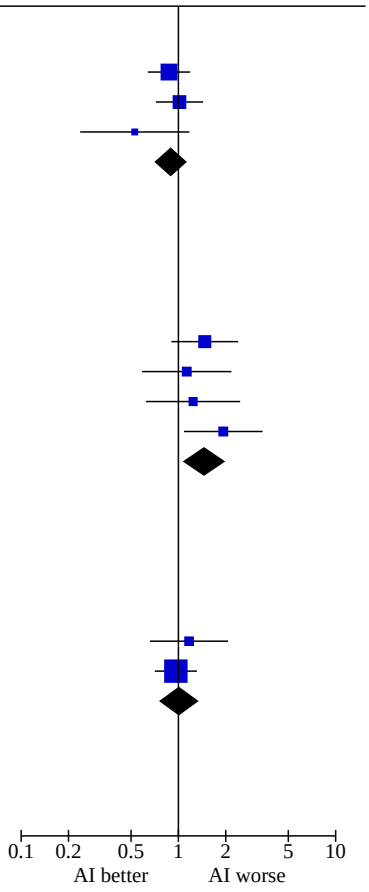
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
4.2.2 AI versus megestrol acetate	4	2036	Odds Ratio (M-H, Fixed, 95% CI)	1.45 [1.09, 1.95]
4.2.3 AI versus fulvestrant	2	1539	Odds Ratio (M-H, Fixed, 95% CI)	1.01 [0.77, 1.32]
4.3 vomiting	5		Odds Ratio (M-H, Fixed, 95% CI)	Subtotals only
4.3.1 AI versus tamoxifen	1	1017	Odds Ratio (M-H, Fixed, 95% CI)	1.07 [0.67, 1.72]
4.3.2 AI versus megestrol acetate	3	1636	Odds Ratio (M-H, Fixed, 95% CI)	1.77 [1.11, 2.83]
4.3.3 AI versus fulvestrant	1	846	Odds Ratio (M-H, Fixed, 95% CI)	0.90 [0.60, 1.35]
4.4 diarrhoea	7		Odds Ratio (M-H, Fixed, 95% CI)	Subtotals only
4.4.1 AI versus tamoxifen	2	1927	Odds Ratio (M-H, Fixed, 95% CI)	1.49 [0.95, 2.35]
4.4.2 AI versus megestrol acetate	3	1278	Odds Ratio (M-H, Fixed, 95% CI)	2.40 [1.34, 4.29]
4.4.3 AI versus fulvestrant	2	1090	Odds Ratio (M-H, Fixed, 95% CI)	1.23 [0.79, 1.90]
4.5 rash	4		Odds Ratio (M-H, Random, 95% CI)	Subtotals only
4.5.1 AI versus megestrol acetate	3	1636	Odds Ratio (M-H, Random, 95% CI)	1.63 [0.47, 5.70]
4.5.2 AI versus fulvestrant	1	397	Odds Ratio (M-H, Random, 95% CI)	1.39 [0.77, 2.50]
4.6 vaginal bleeding	3		Odds Ratio (M-H, Fixed, 95% CI)	Subtotals only
4.6.1 AI versus tamoxifen	1	1017	Odds Ratio (M-H, Fixed, 95% CI)	0.45 [0.16, 1.32]
4.6.2 AI versus megestrol acetate	2	915	Odds Ratio (M-H, Fixed, 95% CI)	0.29 [0.13, 0.65]
4.7 thromboembolic	3		Odds Ratio (M-H, Fixed, 95% CI)	Subtotals only
4.7.1 AI versus tamoxifen	1	1017	Odds Ratio (M-H, Fixed, 95% CI)	0.53 [0.30, 0.96]
4.7.2 AI versus megestrol acetate	1	515	Odds Ratio (M-H, Fixed, 95% CI)	0.71 [0.30, 1.73]
4.7.3 AI versus fulvestrant	1	846	Odds Ratio (M-H, Fixed, 95% CI)	1.14 [0.56, 2.31]
4.8 arthralgia	3		Odds Ratio (M-H, Fixed, 95% CI)	Subtotals only
4.8.1 AI versus tamoxifen	2	1031	Odds Ratio (M-H, Fixed, 95% CI)	1.14 [0.81, 1.60]
4.8.2 AI versus megestrol acetate	1	363	Odds Ratio (M-H, Fixed, 95% CI)	1.77 [0.89, 3.51]

Analysis 4.1. Comparison 4: Current AIs versus non-AI: Toxicity, Outcome 1: hot flushes



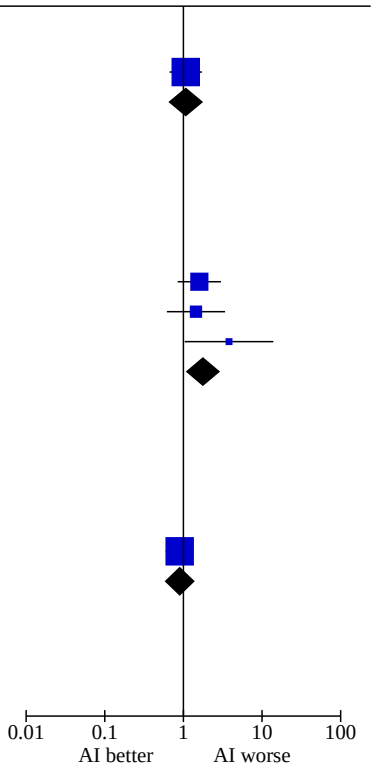
Analysis 4.2. Comparison 4: Current AIs versus non-AI: Toxicity, Outcome 2: nausea

Study or Subgroup	AI Events	AI Total	comparison Events	comparison Total	Weight	Odds Ratio M-H, Fixed, 95% CI
4.2.1 AI versus tamoxifen						
Bonneterre 2001	94	506	106	511	51.6%	0.87 [0.64 , 1.19]
Mourisden 2001	78	455	77	455	38.4%	1.02 [0.72 , 1.44]
Paridaens 2003	14	62	21	59	10.0%	0.53 [0.24 , 1.17]
Subtotal (95% CI)		**1023**		**1025**	**100.0%**	**0.89 [0.72 , 1.11]**
Total events:	186		204			
Heterogeneity: Chi² = 2.22, df = 2 (P = 0.33); I² = 10%						
Test for overall effect: Z = 1.01 (P = 0.31)						
4.2.2 AI versus megestrol acetate						
Buzdar 1996a	46	262	32	253	35.6%	1.47 [0.90 , 2.40]
Buzdar 2001	21	199	19	201	22.4%	1.13 [0.59 , 2.17]
Dombernowsky 1998	19	174	17	189	19.2%	1.24 [0.62 , 2.47]
Kaufmann 2000	33	358	20	400	22.7%	1.93 [1.09 , 3.43]
Subtotal (95% CI)		**993**		**1043**	**100.0%**	**1.45 [1.09 , 1.95]**
Total events:	119		88			
Heterogeneity: Chi² = 1.71, df = 3 (P = 0.64); I² = 0%						
Test for overall effect: Z = 2.52 (P = 0.01)						
4.2.3 AI versus fulvestrant						
Chia 2008	27	342	24	351	21.0%	1.17 [0.66 , 2.07]
Mauriac 2003	107	423	110	423	79.0%	0.96 [0.71 , 1.31]
Subtotal (95% CI)		**765**		**774**	**100.0%**	**1.01 [0.77 , 1.32]**
Total events:	134		134			
Heterogeneity: Chi² = 0.34, df = 1 (P = 0.56); I² = 0%						
Test for overall effect: Z = 0.05 (P = 0.96)						



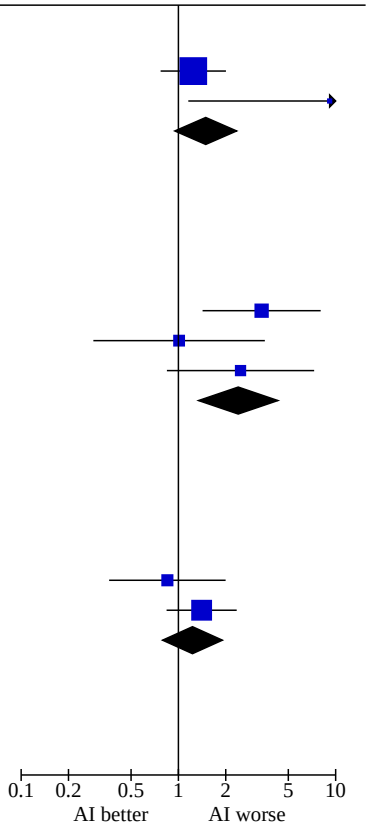
Analysis 4.3. Comparison 4: Current AIs versus non-AI: Toxicity, Outcome 3: vomiting

Study or Subgroup	AI Events	AI Total	non-AI Events	non-AI Total	Weight	Odds Ratio M-H, Fixed, 95% CI
4.3.1 AI versus tamoxifen						
Bonneterre 2001	38	506	36	511	100.0%	1.07 [0.67 , 1.72]
Subtotal (95% CI)		**506**		**511**	**100.0%**	**1.07 [0.67 , 1.72]**
Total events:	38		36			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.29 (P = 0.78)						
4.3.2 AI versus megestrol acetate						
Buzdar 1996a	27	262	17	253	57.2%	1.59 [0.85 , 3.00]
Dombernowsky 1998	13	174	10	189	32.7%	1.45 [0.62 , 3.39]
Kaufmann 2000	10	358	3	400	10.1%	3.80 [1.04 , 13.93]
Subtotal (95% CI)		**794**		**842**	**100.0%**	**1.77 [1.11 , 2.83]**
Total events:	50		30			
Heterogeneity: $Chi^2 = 1.65$, df = 2 (P = 0.44); $I^2 = 0\%$						
Test for overall effect: Z = 2.39 (P = 0.02)						
4.3.3 AI versus fulvestrant						
Mauriac 2003	50	423	55	423	100.0%	0.90 [0.60 , 1.35]
Subtotal (95% CI)		**423**		**423**	**100.0%**	**0.90 [0.60 , 1.35]**
Total events:	50		55			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.52 (P = 0.60)						



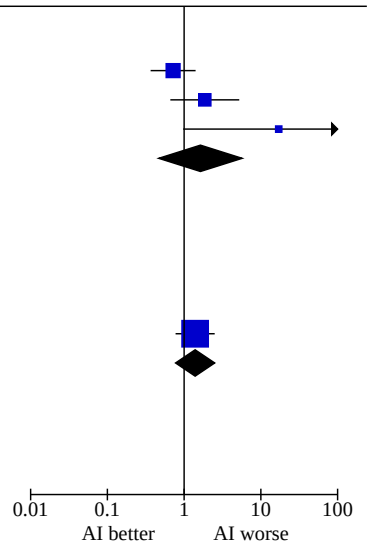
Analysis 4.4. Comparison 4: Current AIs versus non-AI: Toxicity, Outcome 4: diarrhoea

Study or Subgroup	AI Events	AI Total	comparison Events	comparison Total	Weight	Odds Ratio M-H, Fixed, 95% CI
4.4.1 AI versus tamoxifen						
Bonneterre 2001	40	506	33	511	96.9%	1.24 [0.77 , 2.01]
Mourisden 2001	9	455	1	455	3.1%	9.16 [1.16 , 72.61]
Subtotal (95% CI)		**961**		**966**	**100.0%**	**1.49 [0.95 , 2.35]**
Total events:	49		34			
Heterogeneity: Chi² = 3.51, df = 1 (P = 0.06); I² = 72%						
Test for overall effect: Z = 1.73 (P = 0.08)						
4.4.2 AI versus megestrol acetate						
Buzdar 1996a	23	262	7	253	41.0%	3.38 [1.42 , 8.03]
Buzdar 2001	5	199	5	201	30.6%	1.01 [0.29 , 3.55]
Dombernowsky 1998	11	174	5	189	28.4%	2.48 [0.85 , 7.30]
Subtotal (95% CI)		**635**		**643**	**100.0%**	**2.40 [1.34 , 4.29]**
Total events:	39		17			
Heterogeneity: Chi² = 2.43, df = 2 (P = 0.30); I² = 18%						
Test for overall effect: Z = 2.96 (P = 0.003)						
4.4.3 AI versus fulvestrant						
Chia 2008	10	342	12	351	31.8%	0.85 [0.36 , 2.00]
Mauriac 2003	40	193	32	204	68.2%	1.41 [0.84 , 2.35]
Subtotal (95% CI)		**535**		**555**	**100.0%**	**1.23 [0.79 , 1.90]**
Total events:	50		44			
Heterogeneity: Chi² = 0.98, df = 1 (P = 0.32); I² = 0%						
Test for overall effect: Z = 0.92 (P = 0.36)						



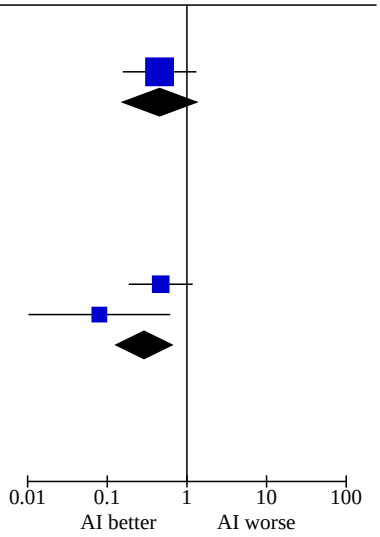
Analysis 4.5. Comparison 4: Current AIs versus non-AI: Toxicity, Outcome 5: rash

Study or Subgroup	comparison Events	comparison Total	AI Events	AI Total	Weight	Odds Ratio M-H, Random, 95% CI
4.5.1 AI versus megestrol acetate						
Buzdar 1996a	16	262	21	253	46.6%	0.72 [0.37 , 1.41]
Dombernowsky 1998	10	174	6	189	39.4%	1.86 [0.66 , 5.23]
Kaufmann 2000	7	358	0	400	14.1%	17.09 [0.97 , 300.32]
Subtotal (95% CI)		**794**		**842**	**100.0%**	**1.63 [0.47 , 5.70]**
Total events:	33		27			
Heterogeneity: Tau² = 0.76; Chi² = 6.35, df = 2 (P = 0.04); I² = 68%						
Test for overall effect: Z = 0.77 (P = 0.44)						
4.5.2 AI versus fulvestrant						
Mauriac 2003	29	193	23	204	100.0%	1.39 [0.77 , 2.50]
Subtotal (95% CI)		**193**		**204**	**100.0%**	**1.39 [0.77 , 2.50]**
Total events:	29		23			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.10 (P = 0.27)						



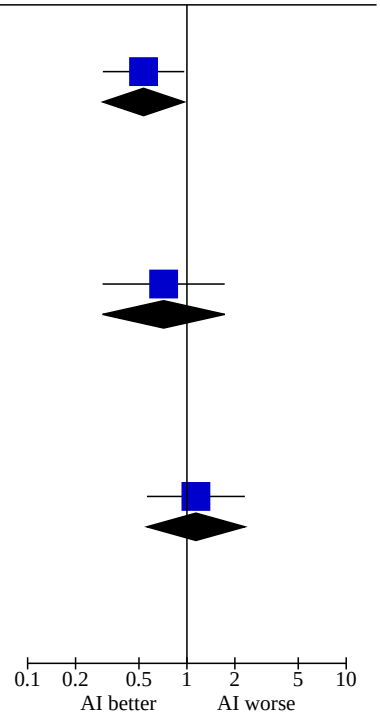
Analysis 4.6. Comparison 4: Current AIs versus non-AI: Toxicity, Outcome 6: vaginal bleeding

Study or Subgroup	AI Events	AI Total	comparison Events	comparison Total	Weight	Odds Ratio M-H, Fixed, 95% CI
4.6.1 AI versus tamoxifen						
Bonneterre 2001	5	506	11	511	100.0%	0.45 [0.16 , 1.32]
Subtotal (95% CI)		**506**		**511**	**100.0%**	**0.45 [0.16 , 1.32]**
Total events:	5		11			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.46 (P = 0.15)						
4.6.2 AI versus megestrol acetate						
Buzdar 1996a	7	262	14	253	53.9%	0.47 [0.19 , 1.18]
Buzdar 2001	1	199	12	201	46.1%	0.08 [0.01 , 0.62]
Subtotal (95% CI)		**461**		**454**	**100.0%**	**0.29 [0.13 , 0.65]**
Total events:	8		26			
Heterogeneity: Chi² = 2.57, df = 1 (P = 0.11); I² = 61%						
Test for overall effect: Z = 3.02 (P = 0.003)						

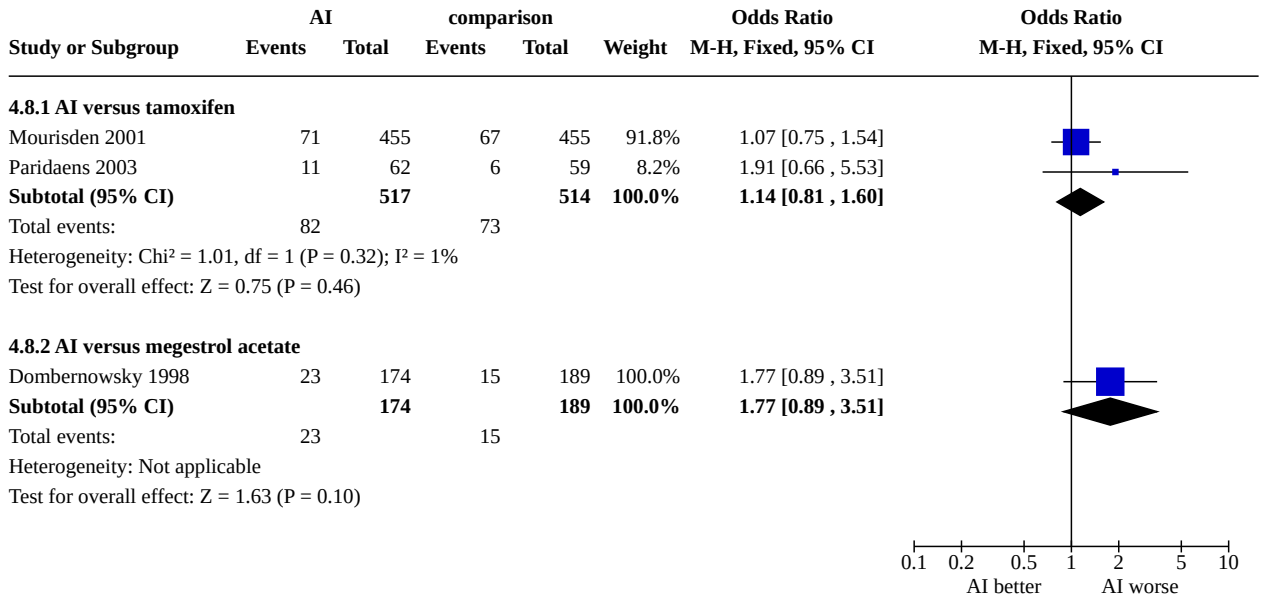


Analysis 4.7. Comparison 4: Current AIs versus non-AI: Toxicity, Outcome 7: thromboembolic

Study or Subgroup	AI Events	AI Total	comparison Events	comparison Total	Weight	Odds Ratio M-H, Fixed, 95% CI
4.7.1 AI versus tamoxifen						
Bonneterre 2001	18	506	33	511	100.0%	0.53 [0.30 , 0.96]
Subtotal (95% CI)		**506**		**511**	**100.0%**	**0.53 [0.30 , 0.96]**
Total events:	18		33			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.09 (P = 0.04)						
4.7.2 AI versus megestrol acetate						
Buzdar 1996a	9	262	12	253	100.0%	0.71 [0.30 , 1.73]
Subtotal (95% CI)		**262**		**253**	**100.0%**	**0.71 [0.30 , 1.73]**
Total events:	9		12			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.75 (P = 0.45)						
4.7.3 AI versus fulvestrant						
Mauriac 2003	17	423	15	423	100.0%	1.14 [0.56 , 2.31]
Subtotal (95% CI)		**423**		**423**	**100.0%**	**1.14 [0.56 , 2.31]**
Total events:	17		15			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.36 (P = 0.72)						



Analysis 4.8. Comparison 4: Current AIs versus non-AI: Toxicity, Outcome 8: arthralgia



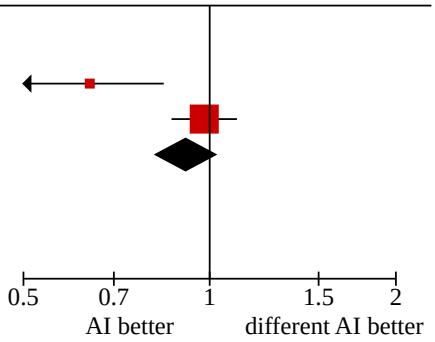
Comparison 5. AI versus different AI

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
5.1 Overall survival (reported)	2		HR (IV, Fixed, 95% CI)	Subtotals only
5.1.1 letrozole	2		HR (IV, Fixed, 95% CI)	0.91 [0.82, 1.02]
5.2 Progession-free survival (reported or calculated)	2		HR (IV, Fixed, 95% CI)	Subtotals only
5.2.1 letrozole	2		HR (IV, Fixed, 95% CI)	0.97 [0.90, 1.04]
5.3 Clinical benefit (assessable)	5		Odds Ratio (M-H, Fixed, 95% CI)	Subtotals only
5.3.1 letrozole	4	1687	Odds Ratio (M-H, Fixed, 95% CI)	0.77 [0.62, 0.95]
5.3.2 anastrozole	2	663	Odds Ratio (M-H, Fixed, 95% CI)	1.29 [0.92, 1.79]
5.4 Objective response (assessable)	5		Odds Ratio (M-H, Fixed, 95% CI)	Subtotals only
5.4.1 letrozole	4	1687	Odds Ratio (M-H, Fixed, 95% CI)	0.62 [0.50, 0.78]
5.4.2 anastrozole	2	663	Odds Ratio (M-H, Fixed, 95% CI)	1.59 [1.07, 2.37]
5.5 Clinical benefit (randomised)	5		Odds Ratio (M-H, Fixed, 95% CI)	Subtotals only
5.5.1 letrozole	4	2098	Odds Ratio (M-H, Fixed, 95% CI)	0.82 [0.68, 0.98]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
5.5.2 anastrozole	2	773	Odds Ratio (M-H, Fixed, 95% CI)	1.25 [0.90, 1.72]
5.6 Objective response (randomised)	5		Odds Ratio (M-H, Fixed, 95% CI)	Subtotals only
5.6.1 letrozole	4	2098	Odds Ratio (M-H, Fixed, 95% CI)	0.66 [0.54, 0.82]
5.6.2 anastrozole	2	782	Odds Ratio (M-H, Fixed, 95% CI)	1.50 [1.01, 2.23]

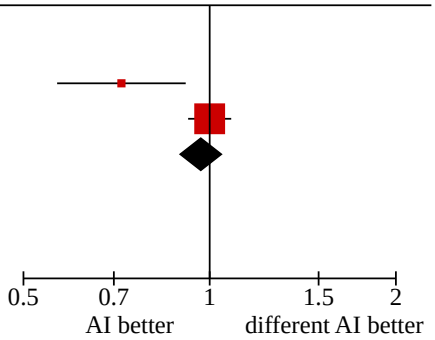
Analysis 5.1. Comparison 5: AI versus different AI, Outcome 1: Overall survival (reported)

Study or Subgroup	log[HR]	SE	Weight	HR IV, Fixed, 95% CI
5.1.1 letrozole				
Gershanovich 1998	-0.4463	0.1405	16.3%	0.64 [0.49 , 0.84]
Goss 2007	-0.0202	0.0621	83.7%	0.98 [0.87 , 1.11]
Subtotal (95% CI)			**100.0%**	**0.91 [0.82 , 1.02]**
Heterogeneity: Chi² = 7.69, df = 1 (P = 0.006); I² = 87%				
Test for overall effect: Z = 1.58 (P = 0.11)				
Test for subgroup differences: Not applicable				



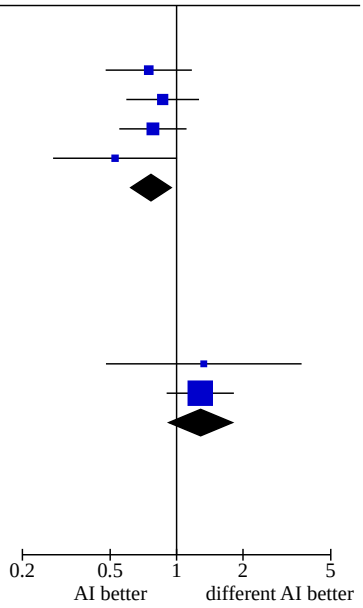
Analysis 5.2. Comparison 5: AI versus different AI, Outcome 2: Progession-free survival (reported or calculated)

Study or Subgroup	log[HR]	SE	Weight	HR IV, Fixed, 95% CI
5.2.1 letrozole				
Gershanovich 1998	-0.3285	0.1221	10.1%	0.72 [0.57 , 0.91]
Goss 2007	0	0.0409	89.9%	1.00 [0.92 , 1.08]
Subtotal (95% CI)			**100.0%**	**0.97 [0.90 , 1.04]**
Heterogeneity: Chi² = 6.51, df = 1 (P = 0.01); I² = 85%				
Test for overall effect: Z = 0.85 (P = 0.39)				
Test for subgroup differences: Not applicable				



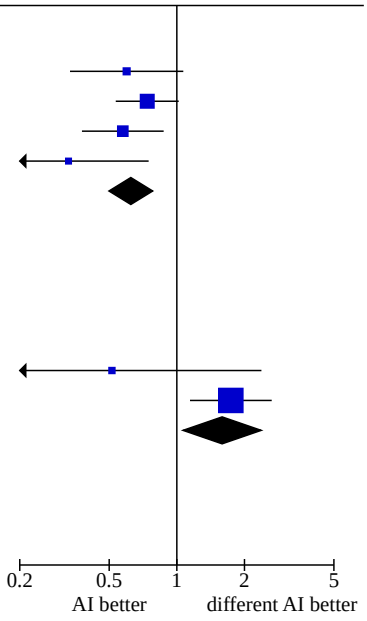
Analysis 5.3. Comparison 5: AI versus different AI, Outcome 3: Clinical benefit (assessable)

Study or Subgroup	different AI better Events	Total	AI better Events	Total	Weight	Odds Ratio M-H, Fixed, 95% CI
5.3.1 letrozole						
Gershanovich 1998	52	162	67	173	22.3%	0.75 [0.48 , 1.17]
Goss 2007	224	298	231	297	29.1%	0.86 [0.59 , 1.26]
Rose 2003	82	304	96	299	35.8%	0.78 [0.55 , 1.11]
Tominaga 2003	27	77	39	77	12.8%	0.53 [0.28 , 1.00]
Subtotal (95% CI)		**841**		**846**	**100.0%**	**0.77 [0.62 , 0.95]**
Total events:	385		433			
Heterogeneity: Chi² = 1.71, df = 3 (P = 0.63); I² = 0%						
Test for overall effect: Z = 2.48 (P = 0.01)						
5.3.2 anastrozole						
Kleeberg 1997	15	31	12	29	10.4%	1.33 [0.48 , 3.69]
Rose 2003	96	299	82	304	89.6%	1.28 [0.90 , 1.82]
Subtotal (95% CI)		**330**		**333**	**100.0%**	**1.29 [0.92 , 1.79]**
Total events:	111		94			
Heterogeneity: Chi² = 0.00, df = 1 (P = 0.95); I² = 0%						
Test for overall effect: Z = 1.48 (P = 0.14)						



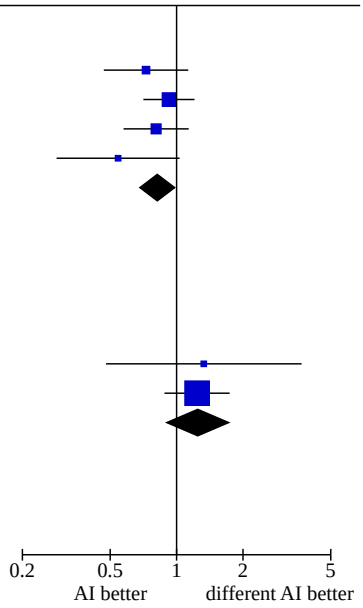
Analysis 5.4. Comparison 5: AI versus different AI, Outcome 4: Objective response (assessable)

Study or Subgroup	different AI Events	Total	AI Events	Total	Weight	Odds Ratio M-H, Fixed, 95% CI
5.4.1 letrozole						
Gershanovich 1998	22	162	36	173	15.4%	0.60 [0.33 , 1.07]
Goss 2007	132	298	154	297	43.9%	0.74 [0.53 , 1.02]
Rose 2003	44	304	68	299	30.0%	0.57 [0.38 , 0.87]
Tominaga 2003	10	77	24	77	10.7%	0.33 [0.15 , 0.75]
Subtotal (95% CI)		**841**		**846**	**100.0%**	**0.62 [0.50 , 0.78]**
Total events:	208		282			
Heterogeneity: Chi² = 3.54, df = 3 (P = 0.32); I² = 15%						
Test for overall effect: Z = 4.13 (P < 0.0001)						
5.4.2 anastrozole						
Kleeberg 1997	3	31	5	29	12.2%	0.51 [0.11 , 2.38]
Rose 2003	68	299	44	304	87.8%	1.74 [1.14 , 2.64]
Subtotal (95% CI)		**330**		**333**	**100.0%**	**1.59 [1.07 , 2.37]**
Total events:	71		49			
Heterogeneity: Chi² = 2.26, df = 1 (P = 0.13); I² = 56%						
Test for overall effect: Z = 2.27 (P = 0.02)						



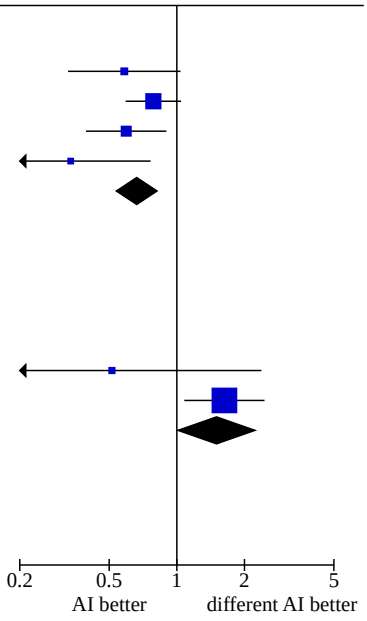
Analysis 5.5. Comparison 5: AI versus different AI, Outcome 5: Clinical benefit (randomised)

Study or Subgroup	different AI better Events	Total	AI better Events	Total	Weight	Odds Ratio M-H, Fixed, 95% CI
5.5.1 letrozole						
Gershanovich 1998	52	178	67	185	18.0%	0.73 [0.47 , 1.13]
Goss 2007	224	434	231	431	43.5%	0.92 [0.71 , 1.21]
Rose 2003	82	357	96	356	28.7%	0.81 [0.57 , 1.13]
Tominaga 2003	27	78	39	79	9.8%	0.54 [0.29 , 1.03]
Subtotal (95% CI)		**1047**		**1051**	**100.0%**	**0.82 [0.68 , 0.98]**
Total events:	385		433			
Heterogeneity: Chi² = 2.64, df = 3 (P = 0.45); I² = 0%						
Test for overall effect: Z = 2.18 (P = 0.03)						
5.5.2 anastrozole						
Kleeberg 1997	15	31	12	29	9.7%	1.33 [0.48 , 3.69]
Rose 2003	96	356	82	357	90.3%	1.24 [0.88 , 1.74]
Subtotal (95% CI)		**387**		**386**	**100.0%**	**1.25 [0.90 , 1.72]**
Total events:	111		94			
Heterogeneity: Chi² = 0.02, df = 1 (P = 0.90); I² = 0%						
Test for overall effect: Z = 1.34 (P = 0.18)						



Analysis 5.6. Comparison 5: AI versus different AI, Outcome 6: Objective response (randomised)

Study or Subgroup	different AI Events	Total	AI Events	Total	Weight	Odds Ratio M-H, Fixed, 95% CI
5.6.1 letrozole						
Gershanovich 1998	22	178	36	185	14.1%	0.58 [0.33 , 1.04]
Goss 2007	132	434	154	431	49.1%	0.79 [0.59 , 1.04]
Rose 2003	44	357	68	356	27.3%	0.60 [0.39 , 0.90]
Tominaga 2003	10	78	24	79	9.5%	0.34 [0.15 , 0.76]
Subtotal (95% CI)		**1047**		**1051**	**100.0%**	**0.66 [0.54 , 0.82]**
Total events:	208		282			
Heterogeneity: Chi² = 4.46, df = 3 (P = 0.22); I² = 33%						
Test for overall effect: Z = 3.87 (P = 0.0001)						
5.6.2 anastrozole						
Kleeberg 1997	3	31	5	29	11.4%	0.51 [0.11 , 2.38]
Rose 2003	68	365	44	357	88.6%	1.63 [1.08 , 2.46]
Subtotal (95% CI)		**396**		**386**	**100.0%**	**1.50 [1.01 , 2.23]**
Total events:	71		49			
Heterogeneity: Chi² = 2.03, df = 1 (P = 0.15); I² = 51%						
Test for overall effect: Z = 2.02 (P = 0.04)						



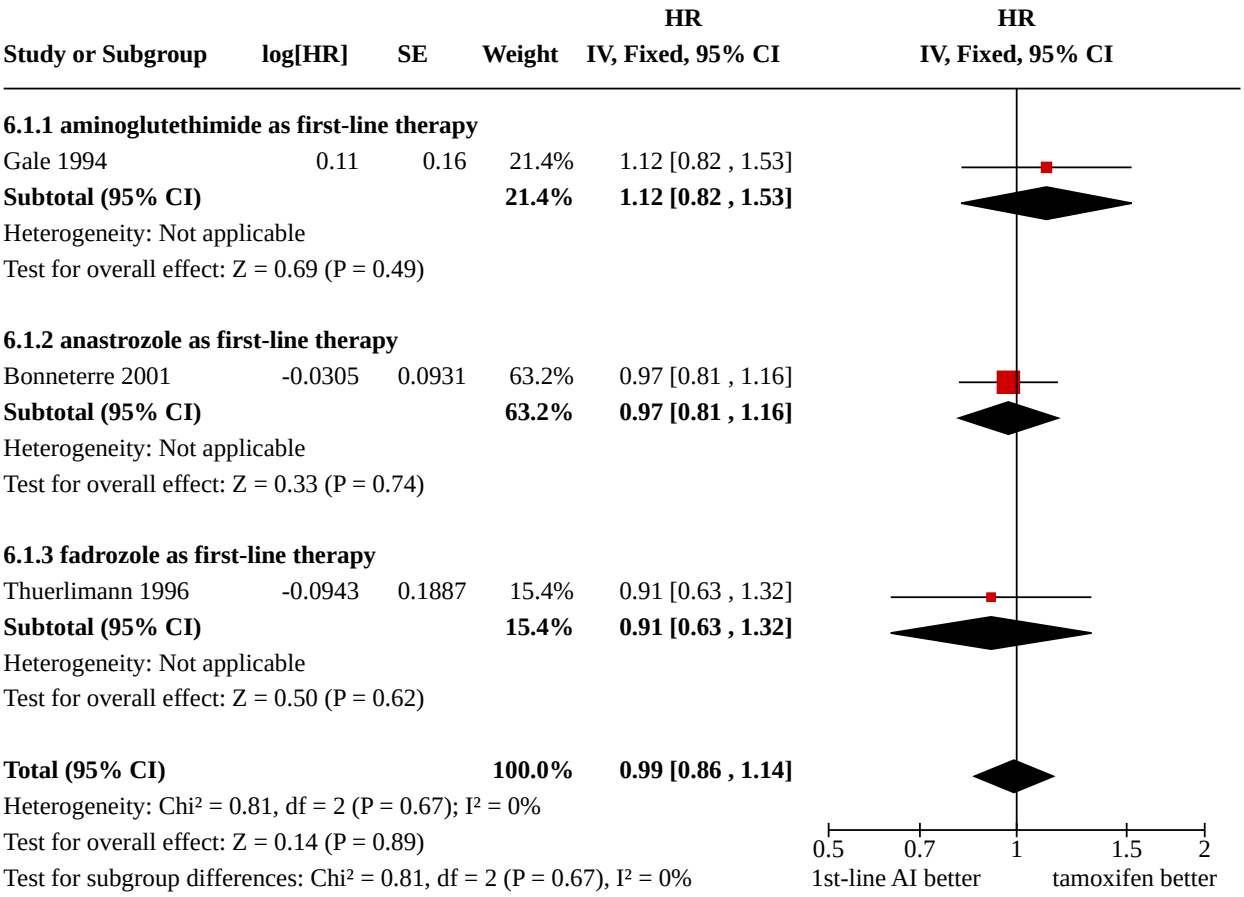
Comparison 6. AI as first-line therapy versus any other therapy (tamoxifen)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
6.1 Overall survival (reported or calculated)	3		HR (IV, Fixed, 95% CI)	0.99 [0.86, 1.14]
6.1.1 aminoglutethimide as first-line therapy	1		HR (IV, Fixed, 95% CI)	1.12 [0.82, 1.53]
6.1.2 anastrozole as first-line therapy	1		HR (IV, Fixed, 95% CI)	0.97 [0.81, 1.16]
6.1.3 fadrozole as first-line therapy	1		HR (IV, Fixed, 95% CI)	0.91 [0.63, 1.32]
6.2 Progression-free survival (reported or calculated)	4		HR (IV, Fixed, 95% CI)	0.78 [0.71, 0.86]
6.2.1 aminoglutethimide	1		HR (IV, Fixed, 95% CI)	0.84 [0.65, 1.08]
6.2.2 formestane as first-line therapy	1		HR (IV, Fixed, 95% CI)	0.93 [0.68, 1.28]
6.2.3 anastrozole as first-line therapy	1		HR (IV, Fixed, 95% CI)	0.82 [0.71, 0.95]
6.2.4 letrozole as first-line therapy	1		HR (IV, Fixed, 95% CI)	0.70 [0.60, 0.82]
6.3 Clinical benefit (assessable)	9	3252	Odds Ratio (M-H, Random, 95% CI)	0.69 [0.51, 0.92]
6.3.1 aminoglutethimide (any dose)	3	479	Odds Ratio (M-H, Random, 95% CI)	0.63 [0.42, 0.93]
6.3.2 formestane 250 mg	1	348	Odds Ratio (M-H, Random, 95% CI)	1.36 [0.87, 2.13]
6.3.3 anastrozole 1 mg	2	1259	Odds Ratio (M-H, Random, 95% CI)	0.48 [0.16, 1.44]
6.3.4 exemestane 25 mg	1	113	Odds Ratio (M-H, Random, 95% CI)	0.47 [0.22, 0.99]
6.3.5 fadrozole 2 mg	1	209	Odds Ratio (M-H, Random, 95% CI)	1.09 [0.58, 2.06]
6.3.6 letrozole 2.5 mg	1	844	Odds Ratio (M-H, Random, 95% CI)	0.63 [0.48, 0.82]
6.4 Objective response (assessable)	11	3503	Odds Ratio (M-H, Random, 95% CI)	0.77 [0.59, 1.00]
6.4.1 aminoglutethimide (any dose)	4	656	Odds Ratio (M-H, Random, 95% CI)	0.75 [0.45, 1.25]

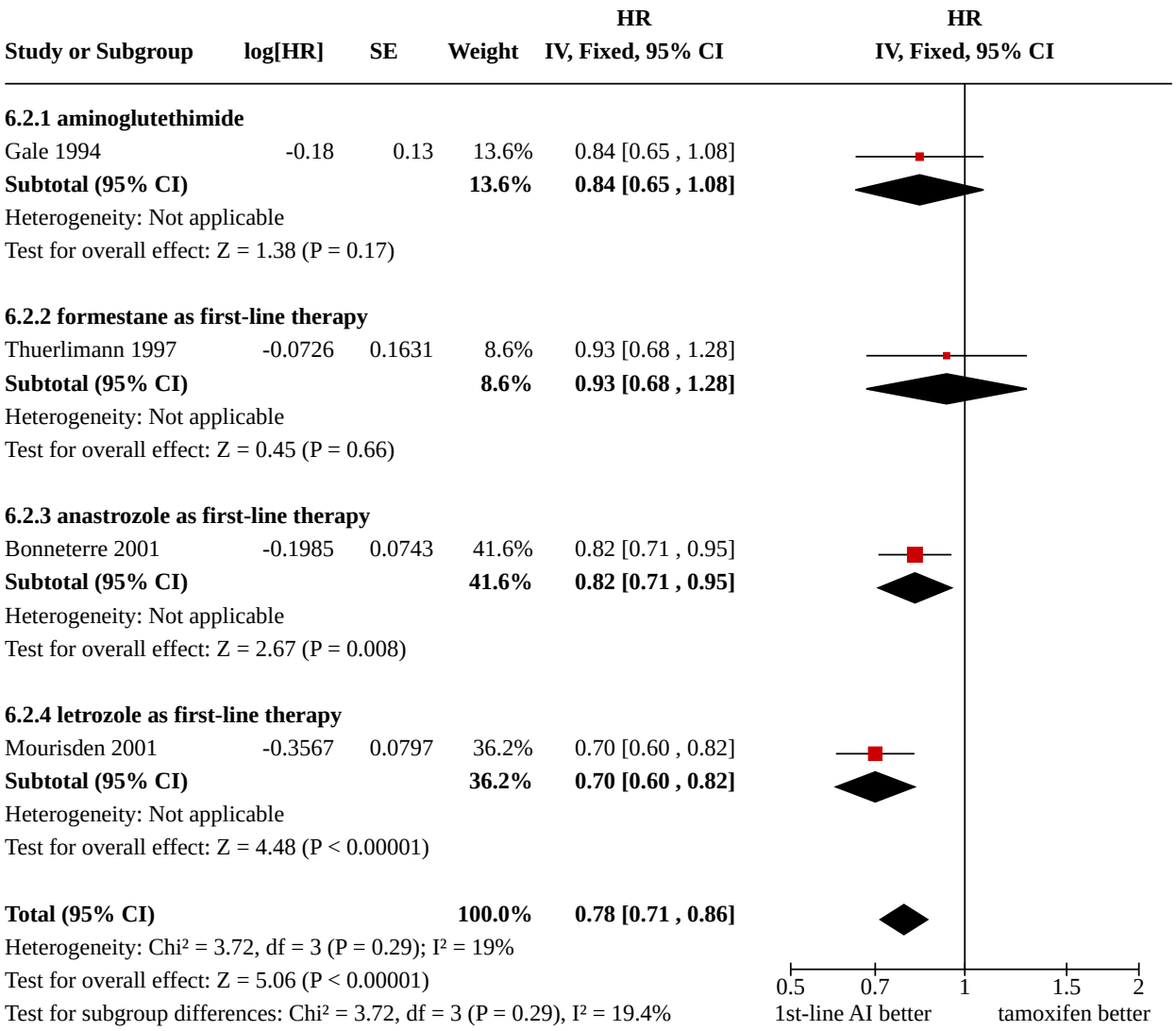
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
6.4.2 formestane 250 mg	1	348	Odds Ratio (M-H, Random, 95% CI)	1.20 [0.77, 1.87]
6.4.3 anastrozole 1 mg	2	1259	Odds Ratio (M-H, Random, 95% CI)	0.85 [0.65, 1.11]
6.4.4 exemestane 25 mg	1	113	Odds Ratio (M-H, Random, 95% CI)	0.26 [0.11, 0.62]
6.4.5 fadrozole 2 mg	2	283	Odds Ratio (M-H, Random, 95% CI)	1.20 [0.69, 2.09]
6.4.6 letrozole 2.5 mg	1	844	Odds Ratio (M-H, Random, 95% CI)	0.58 [0.42, 0.78]
6.5 Clinical benefit (randomised)	9	3451	Odds Ratio (M-H, Random, 95% CI)	0.74 [0.56, 0.98]
6.5.1 aminoglutethimide (any dose)	3	533	Odds Ratio (M-H, Random, 95% CI)	0.74 [0.51, 1.08]
6.5.2 formestane 250 mg	1	409	Odds Ratio (M-H, Random, 95% CI)	1.25 [0.85, 1.86]
6.5.3 anastrozole 1 mg	2	1259	Odds Ratio (M-H, Random, 95% CI)	0.48 [0.16, 1.44]
6.5.4 exemestane 25 mg	1	122	Odds Ratio (M-H, Random, 95% CI)	0.55 [0.27, 1.13]
6.5.5 fadrozole 2 mg	1	221	Odds Ratio (M-H, Random, 95% CI)	1.23 [0.69, 2.21]
6.5.6 letrozole 2.5 mg	1	907	Odds Ratio (M-H, Random, 95% CI)	0.65 [0.50, 0.84]
6.6 Objective response (randomised)	11	3746	Odds Ratio (M-H, Random, 95% CI)	0.81 [0.62, 1.05]
6.6.1 aminoglutethimide (any dose)	4	748	Odds Ratio (M-H, Random, 95% CI)	0.83 [0.48, 1.45]
6.6.2 formestane 250 mg	1	409	Odds Ratio (M-H, Random, 95% CI)	1.18 [0.77, 1.80]
6.6.3 anastrozole 1 mg	2	1259	Odds Ratio (M-H, Random, 95% CI)	0.85 [0.65, 1.11]
6.6.4 exemestane 25 mg	1	122	Odds Ratio (M-H, Random, 95% CI)	0.30 [0.13, 0.69]
6.6.5 fadrozole 2 mg	2	301	Odds Ratio (M-H, Random, 95% CI)	1.28 [0.76, 2.15]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
6.6.6 letrozole 2.5 mg	1	907	Odds Ratio (M-H, Random, 95% CI)	0.59 [0.43, 0.79]

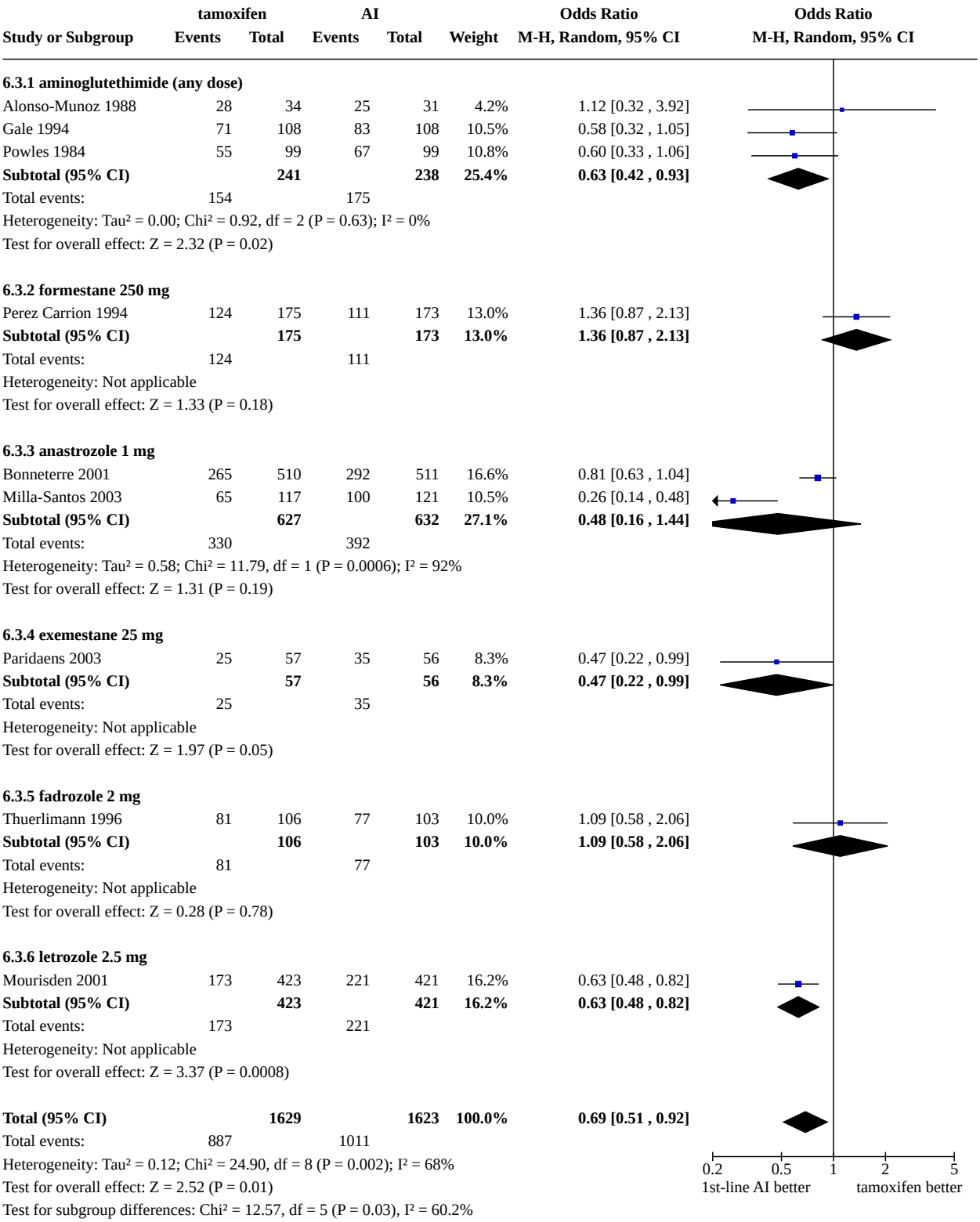
Analysis 6.1. Comparison 6: AI as first-line therapy versus any other therapy (tamoxifen), Outcome 1: Overall survival (reported or calculated)



Analysis 6.2. Comparison 6: AI as first-line therapy versus any other therapy (tamoxifen), Outcome 2: Progression-free survival (reported or calculated)

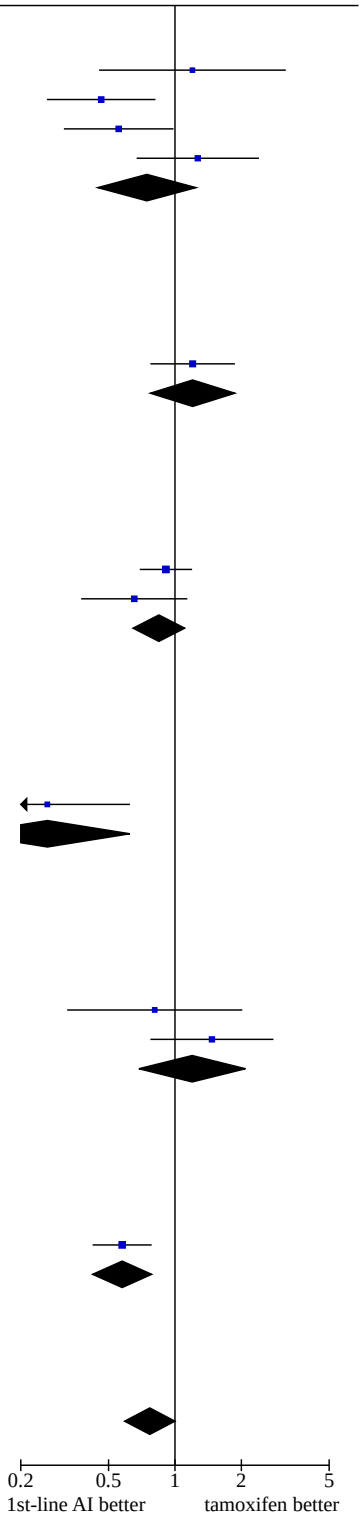


Analysis 6.3. Comparison 6: AI as first-line therapy versus any other therapy (tamoxifen), Outcome 3: Clinical benefit (assessable)

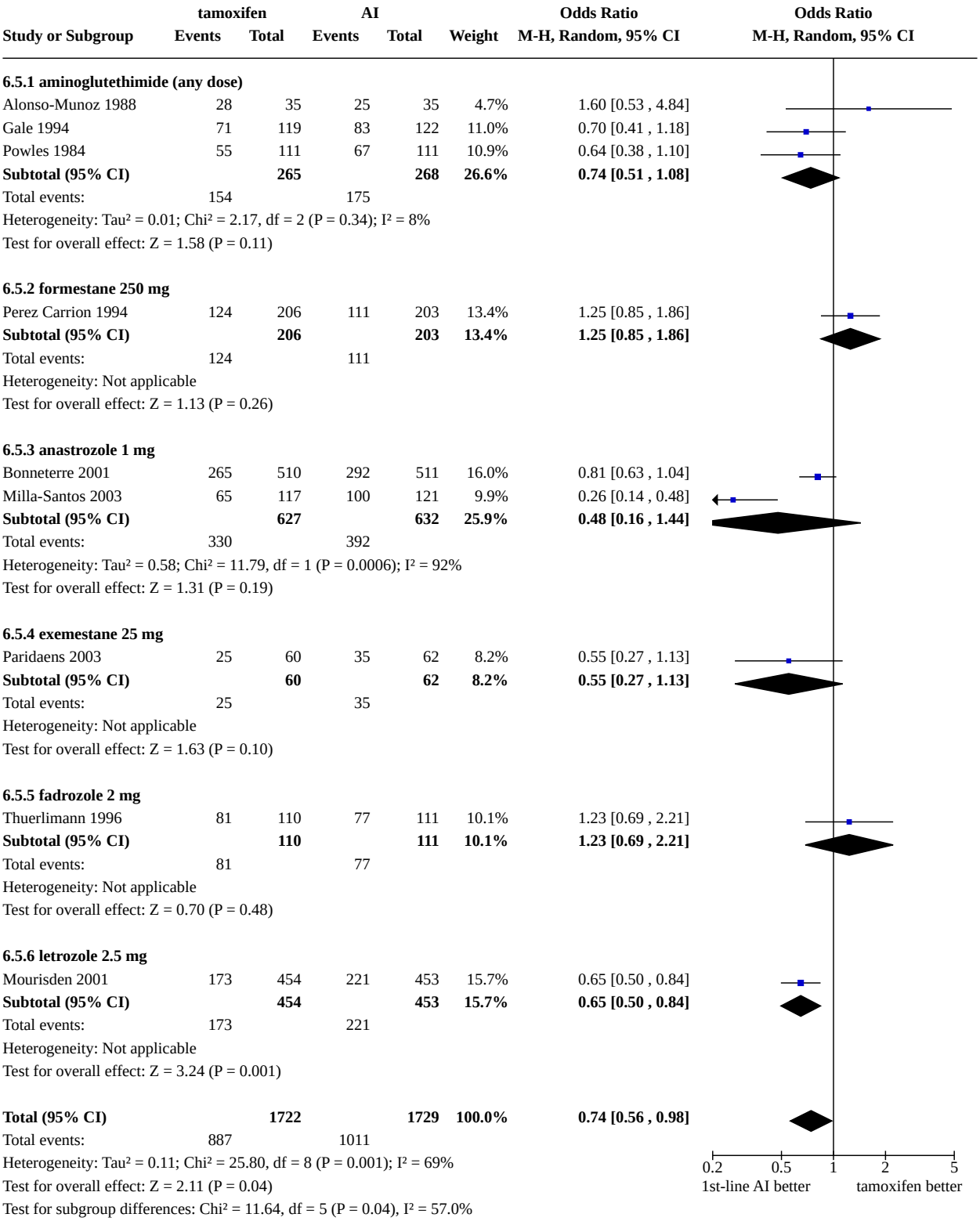


Analysis 6.4. Comparison 6: AI as first-line therapy versus any other therapy (tamoxifen), Outcome 4: Objective response (assessable)

Study or Subgroup	tamoxifen Events	tamoxifen Total	AI Events	AI Total	Weight	Odds Ratio M-H, Random, 95% CI
6.4.1 aminoglutethimide (any dose)						
Alonso-Munoz 1988	18	34	15	31	5.0%	1.20 [0.45 , 3.18]
Gale 1994	30	108	49	108	9.3%	0.46 [0.26 , 0.82]
Powles 1984	34	99	48	99	9.3%	0.56 [0.31 , 0.98]
Rose 1986	32	94	24	83	8.4%	1.27 [0.67 , 2.40]
Subtotal (95% CI)		**335**		**321**	**32.0%**	**0.75 [0.45 , 1.25]**
Total events:	114		136			
Heterogeneity: Tau² = 0.16; Chi² = 7.18, df = 3 (P = 0.07); I² = 58%						
Test for overall effect: Z = 1.12 (P = 0.26)						
6.4.2 formestane 250 mg						
Perez Carrion 1994	65	175	57	173	11.3%	1.20 [0.77 , 1.87]
Subtotal (95% CI)		**175**		**173**	**11.3%**	**1.20 [0.77 , 1.87]**
Total events:	65		57			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.82 (P = 0.41)						
6.4.3 anastrozole 1 mg						
Bonneterre 2001	138	510	148	511	14.0%	0.91 [0.69 , 1.20]
Milla-Santos 2003	31	117	43	121	9.5%	0.65 [0.38 , 1.14]
Subtotal (95% CI)		**627**		**632**	**23.5%**	**0.85 [0.65 , 1.11]**
Total events:	169		191			
Heterogeneity: Tau² = 0.00; Chi² = 1.10, df = 1 (P = 0.29); I² = 9%						
Test for overall effect: Z = 1.22 (P = 0.22)						
6.4.4 exemestane 25 mg						
Paridaens 2003	10	57	25	56	5.9%	0.26 [0.11 , 0.62]
Subtotal (95% CI)		**57**		**56**	**5.9%**	**0.26 [0.11 , 0.62]**
Total events:	10		25			
Heterogeneity: Not applicable						
Test for overall effect: Z = 3.03 (P = 0.002)						
6.4.5 fadrozole 2 mg						
Falkson 1996	17	38	18	36	5.5%	0.81 [0.32 , 2.02]
Thuerlimann 1996	29	106	21	103	8.3%	1.47 [0.77 , 2.79]
Subtotal (95% CI)		**144**		**139**	**13.8%**	**1.20 [0.69 , 2.09]**
Total events:	46		39			
Heterogeneity: Tau² = 0.02; Chi² = 1.10, df = 1 (P = 0.29); I² = 9%						
Test for overall effect: Z = 0.63 (P = 0.53)						
6.4.6 letrozole 2.5 mg						
Mourisden 2001	92	423	137	421	13.4%	0.58 [0.42 , 0.78]
Subtotal (95% CI)		**423**		**421**	**13.4%**	**0.58 [0.42 , 0.78]**
Total events:	92		137			
Heterogeneity: Not applicable						
Test for overall effect: Z = 3.51 (P = 0.0005)						
Total (95% CI)		**1761**		**1742**	**100.0%**	**0.77 [0.59 , 1.00]**
Total events:	496		585			
Heterogeneity: Tau² = 0.11; Chi² = 26.42, df = 10 (P = 0.003); I² = 62%						
Test for overall effect: Z = 1.96 (P = 0.05)						
Test for subgroup differences: Chi² = 16.13, df = 5 (P = 0.006), I² = 69.0%						

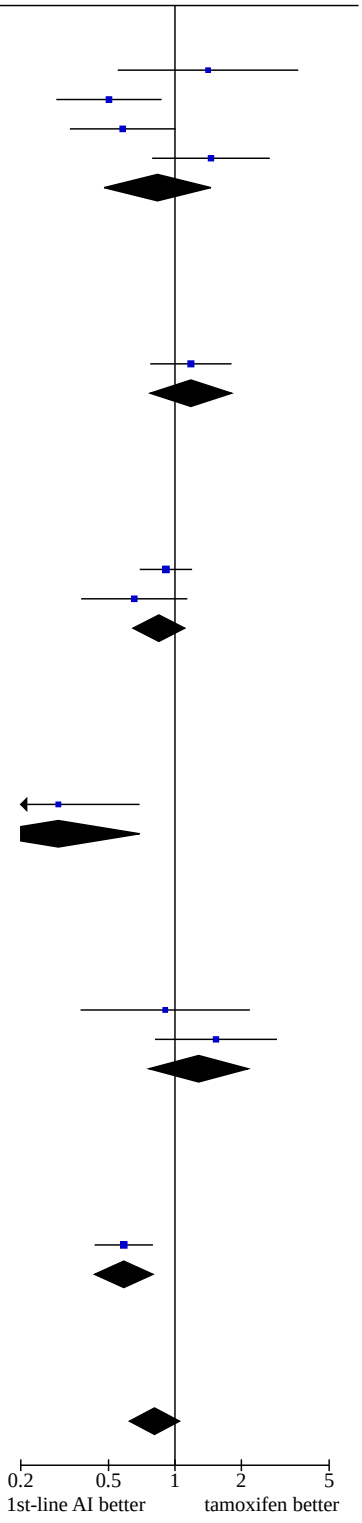


Analysis 6.5. Comparison 6: AI as first-line therapy versus any other therapy (tamoxifen), Outcome 5: Clinical benefit (randomised)



Analysis 6.6. Comparison 6: AI as first-line therapy versus any other therapy (tamoxifen), Outcome 6: Objective response (randomised)

Study or Subgroup	tamoxifen Events	tamoxifen Total	AI Events	AI Total	Weight	Odds Ratio M-H, Random, 95% CI
6.6.1 aminoglutethimide (any dose)						
Alonso-Munoz 1988	18	35	15	35	5.2%	1.41 [0.55 , 3.62]
Gale 1994	30	119	49	122	9.4%	0.50 [0.29 , 0.87]
Powles 1984	34	111	48	111	9.4%	0.58 [0.33 , 1.01]
Rose 1986	32	108	24	107	8.5%	1.46 [0.79 , 2.69]
Subtotal (95% CI)		**373**		**375**	**32.5%**	**0.83 [0.48 , 1.45]**
Total events:	114		136			
Heterogeneity: Tau² = 0.21; Chi² = 9.07, df = 3 (P = 0.03); I² = 67%						
Test for overall effect: Z = 0.64 (P = 0.52)						
6.6.2 formestane 250 mg						
Perez Carrion 1994	65	206	57	203	11.3%	1.18 [0.77 , 1.80]
Subtotal (95% CI)		**206**		**203**	**11.3%**	**1.18 [0.77 , 1.80]**
Total events:	65		57			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.77 (P = 0.44)						
6.6.3 anastrozole 1 mg						
Bonneterre 2001	138	510	148	511	13.7%	0.91 [0.69 , 1.20]
Milla-Santos 2003	31	117	43	121	9.3%	0.65 [0.38 , 1.14]
Subtotal (95% CI)		**627**		**632**	**23.1%**	**0.85 [0.65 , 1.11]**
Total events:	169		191			
Heterogeneity: Tau² = 0.00; Chi² = 1.10, df = 1 (P = 0.29); I² = 9%						
Test for overall effect: Z = 1.22 (P = 0.22)						
6.6.4 exemestane 25 mg						
Paridaens 2003	10	60	25	62	6.0%	0.30 [0.13 , 0.69]
Subtotal (95% CI)		**60**		**62**	**6.0%**	**0.30 [0.13 , 0.69]**
Total events:	10		25			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.82 (P = 0.005)						
6.6.5 fadrozole 2 mg						
Falkson 1996	17	40	18	40	5.7%	0.90 [0.37 , 2.19]
Thuerlimann 1996	29	110	21	111	8.2%	1.53 [0.81 , 2.90]
Subtotal (95% CI)		**150**		**151**	**13.9%**	**1.28 [0.76 , 2.15]**
Total events:	46		39			
Heterogeneity: Tau² = 0.00; Chi² = 0.91, df = 1 (P = 0.34); I² = 0%						
Test for overall effect: Z = 0.94 (P = 0.35)						
6.6.6 letrozole 2.5 mg						
Mourisden 2001	92	454	137	453	13.3%	0.59 [0.43 , 0.79]
Subtotal (95% CI)		**454**		**453**	**13.3%**	**0.59 [0.43 , 0.79]**
Total events:	92		137			
Heterogeneity: Not applicable						
Test for overall effect: Z = 3.44 (P = 0.0006)						
Total (95% CI)		**1870**		**1876**	**100.0%**	**0.81 [0.62 , 1.05]**
Total events:	496		585			
Heterogeneity: Tau² = 0.11; Chi² = 27.11, df = 10 (P = 0.003); I² = 63%						
Test for overall effect: Z = 1.61 (P = 0.11)						
Test for subgroup differences: Chi² = 15.89, df = 5 (P = 0.007), I² = 68.5%						

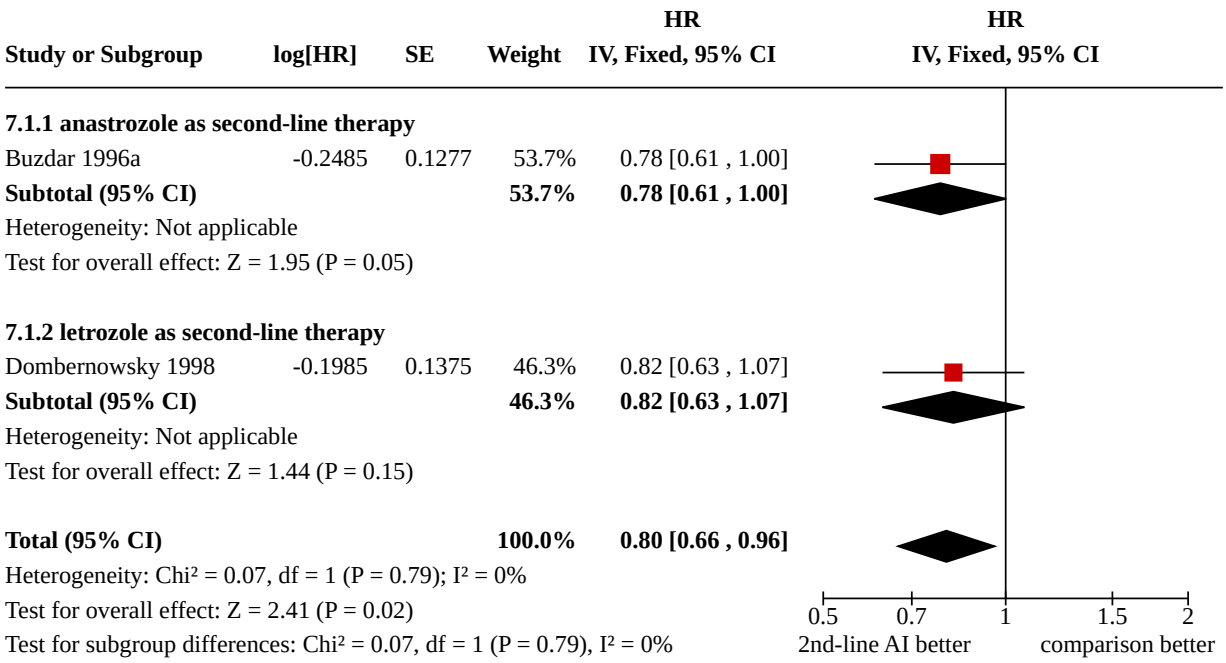


Comparison 7. AI as second-line therapy versus any other therapy

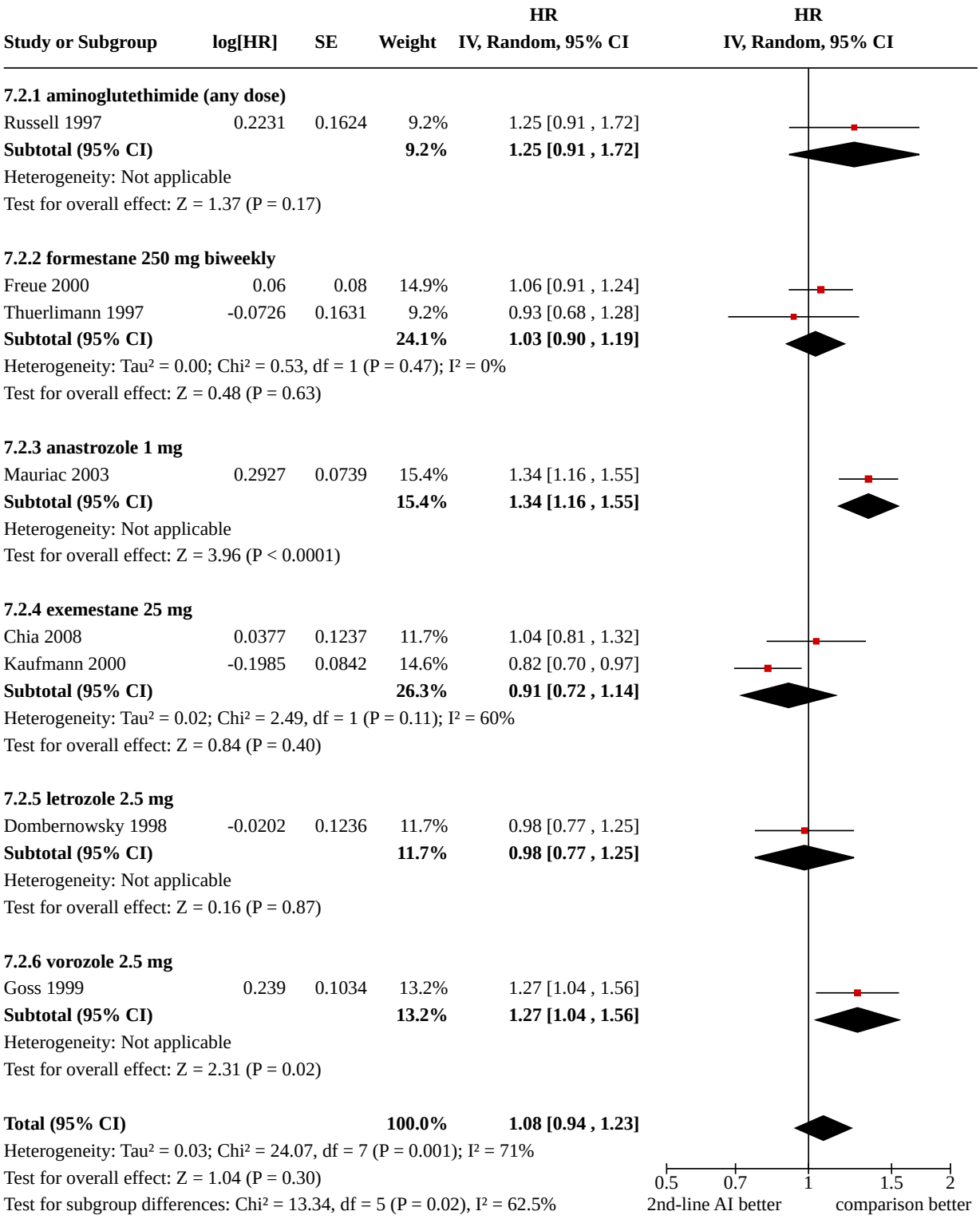
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
7.1 Overall survival (reported or calculated)	2		HR (IV, Fixed, 95% CI)	0.80 [0.66, 0.96]
7.1.1 anastrozole as second-line therapy	1		HR (IV, Fixed, 95% CI)	0.78 [0.61, 1.00]
7.1.2 letrozole as second-line therapy	1		HR (IV, Fixed, 95% CI)	0.82 [0.63, 1.07]
7.2 Progression-free survival (reported or calculated)	8		HR (IV, Random, 95% CI)	1.08 [0.94, 1.23]
7.2.1 aminoglutethimide (any dose)	1		HR (IV, Random, 95% CI)	1.25 [0.91, 1.72]
7.2.2 formestane 250 mg bi-weekly	2		HR (IV, Random, 95% CI)	1.03 [0.90, 1.19]
7.2.3 anastrozole 1 mg	1		HR (IV, Random, 95% CI)	1.34 [1.16, 1.55]
7.2.4 exemestane 25 mg	2		HR (IV, Random, 95% CI)	0.91 [0.72, 1.14]
7.2.5 letrozole 2.5 mg	1		HR (IV, Random, 95% CI)	0.98 [0.77, 1.25]
7.2.6 vorozole 2.5 mg	1		HR (IV, Random, 95% CI)	1.27 [1.04, 1.56]
7.3 Clinical benefit (assessable)	16	5410	Odds Ratio (M-H, Fixed, 95% CI)	0.99 [0.88, 1.11]
7.3.1 aminoglutethimide (any dose)	4	686	Odds Ratio (M-H, Fixed, 95% CI)	0.90 [0.66, 1.23]
7.3.2 formestane 250 mg bi-weekly	1	173	Odds Ratio (M-H, Fixed, 95% CI)	0.75 [0.41, 1.39]
7.3.3 anastrozole 1mg	2	1367	Odds Ratio (M-H, Fixed, 95% CI)	1.04 [0.84, 1.29]
7.3.4 exemestane 25 mg	2	1243	Odds Ratio (M-H, Fixed, 95% CI)	0.95 [0.75, 1.20]
7.3.5 fadrozole 2 mg	3	773	Odds Ratio (M-H, Fixed, 95% CI)	1.04 [0.77, 1.41]
7.3.6 letrozole 2.5 mg	3	793	Odds Ratio (M-H, Fixed, 95% CI)	0.91 [0.68, 1.23]
7.3.7 vorozole 2.5mg	1	375	Odds Ratio (M-H, Fixed, 95% CI)	1.35 [0.88, 2.07]
7.4 Objective response (assessable)	18	5937	Odds Ratio (M-H, Fixed, 95% CI)	0.98 [0.86, 1.13]
7.4.1 aminoglutethimide (any dose)	5	734	Odds Ratio (M-H, Fixed, 95% CI)	0.94 [0.68, 1.30]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
7.4.2 formestane 250 mg bi-weekly	2	652	Odds Ratio (M-H, Fixed, 95% CI)	1.24 [0.84, 1.83]
7.4.3 anastrozole 1 mg	2	1367	Odds Ratio (M-H, Fixed, 95% CI)	1.12 [0.84, 1.50]
7.4.4 exemestane 25 mg	2	1243	Odds Ratio (M-H, Fixed, 95% CI)	0.89 [0.63, 1.26]
7.4.5 fadrozole 2 mg	3	773	Odds Ratio (M-H, Fixed, 95% CI)	1.17 [0.76, 1.80]
7.4.6 letrozole 2.5 mg	3	793	Odds Ratio (M-H, Fixed, 95% CI)	0.76 [0.53, 1.08]
7.4.7 vorozole 2.5 mg	1	375	Odds Ratio (M-H, Fixed, 95% CI)	0.70 [0.34, 1.42]
7.5 Clinical benefit (randomised)	16	6432	Odds Ratio (M-H, Fixed, 95% CI)	1.00 [0.90, 1.11]
7.5.1 aminoglutethimide (any dose)	4	1320	Odds Ratio (M-H, Fixed, 95% CI)	1.06 [0.85, 1.31]
7.5.2 formestane 250 mg bi-weekly	1	177	Odds Ratio (M-H, Fixed, 95% CI)	0.72 [0.40, 1.31]
7.5.3 anastrozole 1 mg	2	1367	Odds Ratio (M-H, Fixed, 95% CI)	1.04 [0.84, 1.29]
7.5.4 exemestane 25 mg	2	1462	Odds Ratio (M-H, Fixed, 95% CI)	0.93 [0.74, 1.16]
7.5.5 fadrozole 2 mg	3	779	Odds Ratio (M-H, Fixed, 95% CI)	1.04 [0.77, 1.41]
7.5.6 letrozole 2.5 mg	3	875	Odds Ratio (M-H, Fixed, 95% CI)	0.91 [0.68, 1.21]
7.5.7 vorozole 2.5mg	1	452	Odds Ratio (M-H, Fixed, 95% CI)	1.25 [0.83, 1.88]
7.6 Objective response (randomised)	18	7113	Odds Ratio (M-H, Fixed, 95% CI)	1.04 [0.92, 1.18]
7.6.1 aminoglutethimide (any dose)	5	1475	Odds Ratio (M-H, Fixed, 95% CI)	1.15 [0.91, 1.45]
7.6.2 formestane 250 mg bi-weekly	2	724	Odds Ratio (M-H, Fixed, 95% CI)	1.22 [0.84, 1.79]
7.6.3 anastrozole 1 mg	2	1367	Odds Ratio (M-H, Fixed, 95% CI)	1.12 [0.84, 1.50]
7.6.4 exemestane 25 mg	2	1462	Odds Ratio (M-H, Fixed, 95% CI)	0.87 [0.62, 1.24]
7.6.5 fadrozole 2 mg	3	779	Odds Ratio (M-H, Fixed, 95% CI)	1.17 [0.76, 1.80]
7.6.6 letrozole 2.5 mg	3	854	Odds Ratio (M-H, Fixed, 95% CI)	0.80 [0.56, 1.13]
7.6.7 vorozole 2.5 mg	1	452	Odds Ratio (M-H, Fixed, 95% CI)	0.67 [0.33, 1.37]

Analysis 7.1. Comparison 7: AI as second-line therapy versus any other therapy, Outcome 1: Overall survival (reported or calculated)

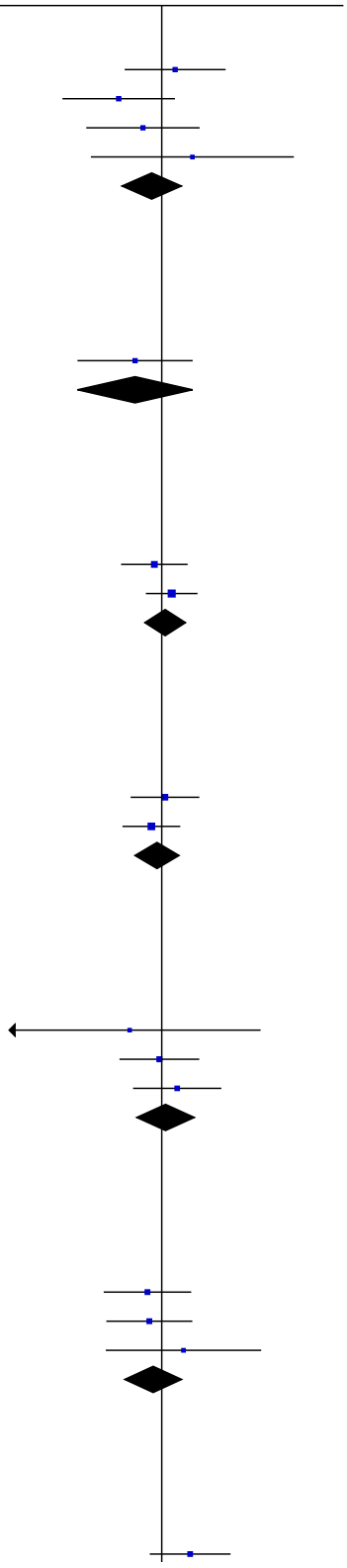


Analysis 7.2. Comparison 7: AI as second-line therapy versus any other therapy, Outcome 2: Progression-free survival (reported or calculated)



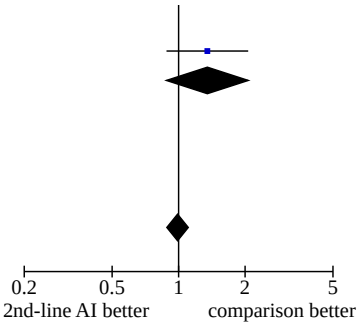
Analysis 7.3. Comparison 7: AI as second-line therapy versus any other therapy, Outcome 3: Clinical benefit (assessable)

Study or Subgroup	comparison Events	Total	2nd line AI Events	Total	Weight	Odds Ratio M-H, Fixed, 95% CI
7.3.1 aminoglutethimide (any dose)						
Canney 1988	61	112	54	106	4.1%	1.15 [0.68 , 1.96]
Garcia-Giralt 1992	79	112	98	124	4.4%	0.64 [0.35 , 1.15]
Lundgren 1989	49	90	51	86	3.8%	0.82 [0.45 , 1.49]
Mercer 1993	13	29	10	27	0.9%	1.38 [0.47 , 4.03]
Subtotal (95% CI)		**343**		**343**	**13.2%**	**0.90 [0.66 , 1.23]**
Total events:	202		213			
Heterogeneity: Chi² = 2.86, df = 3 (P = 0.41); I² = 0%						
Test for overall effect: Z = 0.66 (P = 0.51)						
7.3.2 formestane 250 mg biweekly						
Thuerlimann 1997	46	83	56	90	3.9%	0.75 [0.41 , 1.39]
Subtotal (95% CI)		**83**		**90**	**3.9%**	**0.75 [0.41 , 1.39]**
Total events:	46		56			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.91 (P = 0.36)						
7.3.3 anastrozole 1mg						
Buzdar 1996a	102	253	111	263	10.5%	0.93 [0.65 , 1.31]
Mauriac 2003	186	428	173	423	15.8%	1.11 [0.85 , 1.46]
Subtotal (95% CI)		**681**		**686**	**26.3%**	**1.04 [0.84 , 1.29]**
Total events:	288		284			
Heterogeneity: Chi² = 0.65, df = 1 (P = 0.42); I² = 0%						
Test for overall effect: Z = 0.33 (P = 0.74)						
7.3.4 exemestane 25 mg						
Chia 2008	87	270	85	270	9.3%	1.03 [0.72 , 1.49]
Kaufmann 2000	135	366	133	337	14.1%	0.90 [0.66 , 1.22]
Subtotal (95% CI)		**636**		**607**	**23.3%**	**0.95 [0.75 , 1.20]**
Total events:	222		218			
Heterogeneity: Chi² = 0.35, df = 1 (P = 0.55); I² = 0%						
Test for overall effect: Z = 0.42 (P = 0.67)						
7.3.5 fadrozole 2 mg						
Bezwoda 1998	4	50	5	46	0.8%	0.71 [0.18 , 2.84]
Buzdar 1996b	65	184	70	195	7.1%	0.98 [0.64 , 1.49]
Buzdar 1996c	61	148	56	150	5.3%	1.18 [0.74 , 1.87]
Subtotal (95% CI)		**382**		**391**	**13.1%**	**1.04 [0.77 , 1.41]**
Total events:	130		131			
Heterogeneity: Chi² = 0.65, df = 2 (P = 0.72); I² = 0%						
Test for overall effect: Z = 0.26 (P = 0.80)						
7.3.6 letrozole 2.5 mg						
Buzdar 2001	47	180	53	182	6.3%	0.86 [0.54 , 1.36]
Dombernowsky 1998	60	166	60	153	6.4%	0.88 [0.56 , 1.38]
Schmid 2001	19	60	14	52	1.6%	1.26 [0.55 , 2.85]
Subtotal (95% CI)		**406**		**387**	**14.3%**	**0.91 [0.68 , 1.23]**
Total events:	126		127			
Heterogeneity: Chi² = 0.68, df = 2 (P = 0.71); I² = 0%						
Test for overall effect: Z = 0.59 (P = 0.56)						
7.3.7 vorozole 2.5mg						
Goss 1999	71	185	60	190	5.9%	1.35 [0.88 , 2.07]



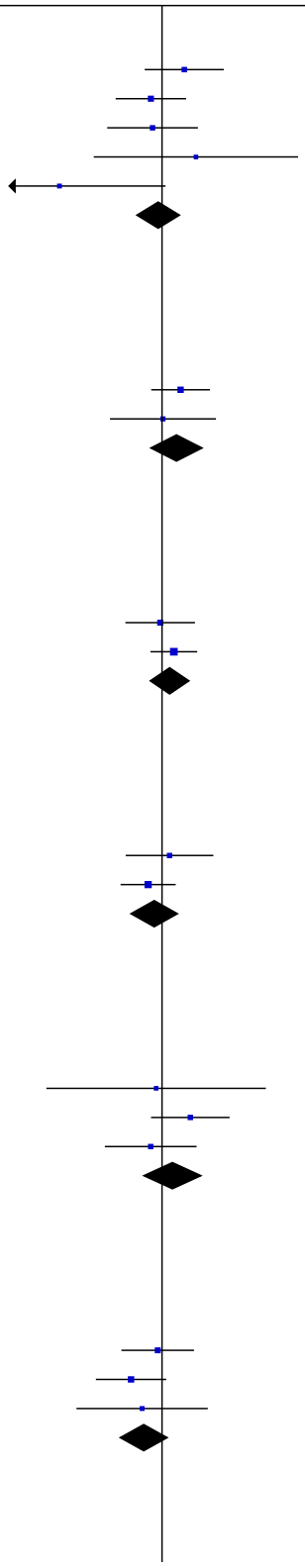
Analysis 7.3. (Continued)

7.3.7 vorozole 2.5mg						
Goss 1999	71	185	60	190	5.9%	1.35 [0.88 , 2.07]
Subtotal (95% CI)		**185**		**190**	**5.9%**	**1.35 [0.88 , 2.07]**
Total events:	71		60			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.38 (P = 0.17)						
Total (95% CI)		**2716**		**2694**	**100.0%**	**0.99 [0.88 , 1.11]**
Total events:	1085		1089			
Heterogeneity: Chi² = 9.02, df = 15 (P = 0.88); I² = 0%						
Test for overall effect: Z = 0.19 (P = 0.85)						
Test for subgroup differences: Chi² = 3.82, df = 6 (P = 0.70), I² = 0%						



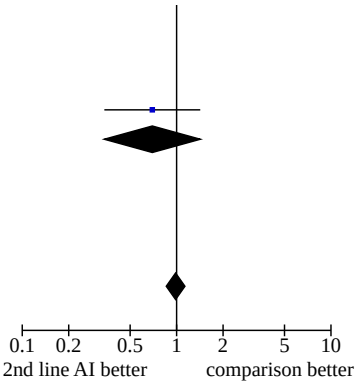
Analysis 7.4. Comparison 7: AI as second-line therapy versus any other therapy, Outcome 4: Objective response (assessable)

Study or Subgroup	comparison Events	Total	2nd line AI Events	Total	Weight	Odds Ratio M-H, Fixed, 95% CI
7.4.1 aminoglutethimide (any dose)						
Canney 1988	35	112	26	106	4.6%	1.40 [0.77 , 2.54]
Garcia-Giralt 1992	39	112	48	124	7.4%	0.85 [0.50 , 1.44]
Lundgren 1989	23	74	26	76	4.4%	0.87 [0.44 , 1.72]
Mercer 1993	5	29	3	27	0.6%	1.67 [0.36 , 7.77]
Russell 1997	2	32	10	42	2.0%	0.21 [0.04 , 1.05]
Subtotal (95% CI)		**359**		**375**	**19.1%**	**0.94 [0.68 , 1.30]**
Total events:	104		113			
Heterogeneity: Chi² = 5.75, df = 4 (P = 0.22); I² = 30%						
Test for overall effect: Z = 0.35 (P = 0.73)						
7.4.2 formestane 250 mg biweekly						
Freue 2000	55	237	45	242	8.5%	1.32 [0.85 , 2.06]
Thuerlimann 1997	14	83	15	90	3.0%	1.01 [0.46 , 2.25]
Subtotal (95% CI)		**320**		**332**	**11.5%**	**1.24 [0.84 , 1.83]**
Total events:	69		60			
Heterogeneity: Chi² = 0.32, df = 1 (P = 0.57); I² = 0%						
Test for overall effect: Z = 1.10 (P = 0.27)						
7.4.3 anastrozole 1 mg						
Buzdar 1996a	31	253	33	263	7.1%	0.97 [0.58 , 1.64]
Mauriac 2003	82	428	70	423	14.2%	1.20 [0.84 , 1.70]
Subtotal (95% CI)		**681**		**686**	**21.3%**	**1.12 [0.84 , 1.50]**
Total events:	113		103			
Heterogeneity: Chi² = 0.41, df = 1 (P = 0.52); I² = 0%						
Test for overall effect: Z = 0.77 (P = 0.44)						
7.4.4 exemestane 25 mg						
Chia 2008	20	270	18	270	4.2%	1.12 [0.58 , 2.17]
Kaufmann 2000	50	366	55	337	12.3%	0.81 [0.54 , 1.23]
Subtotal (95% CI)		**636**		**607**	**16.5%**	**0.89 [0.63 , 1.26]**
Total events:	70		73			
Heterogeneity: Chi² = 0.66, df = 1 (P = 0.42); I² = 0%						
Test for overall effect: Z = 0.66 (P = 0.51)						
7.4.5 fadrozole 2 mg						
Bezwoda 1998	3	50	3	46	0.7%	0.91 [0.18 , 4.78]
Buzdar 1996b	30	184	22	195	4.5%	1.53 [0.85 , 2.77]
Buzdar 1996c	17	148	20	150	4.4%	0.84 [0.42 , 1.68]
Subtotal (95% CI)		**382**		**391**	**9.6%**	**1.17 [0.76 , 1.80]**
Total events:	50		45			
Heterogeneity: Chi² = 1.74, df = 2 (P = 0.42); I² = 0%						
Test for overall effect: Z = 0.71 (P = 0.48)						
7.4.6 letrozole 2.5 mg						
Buzdar 2001	30	180	32	182	6.6%	0.94 [0.54 , 1.62]
Dombernowsky 1998	31	166	41	153	8.7%	0.63 [0.37 , 1.07]
Schmid 2001	9	60	10	52	2.3%	0.74 [0.28 , 1.99]
Subtotal (95% CI)		**406**		**387**	**17.5%**	**0.76 [0.53 , 1.08]**
Total events:	70		83			
Heterogeneity: Chi² = 1.07, df = 2 (P = 0.58); I² = 0%						
Test for overall effect: Z = 1.53 (P = 0.13)						



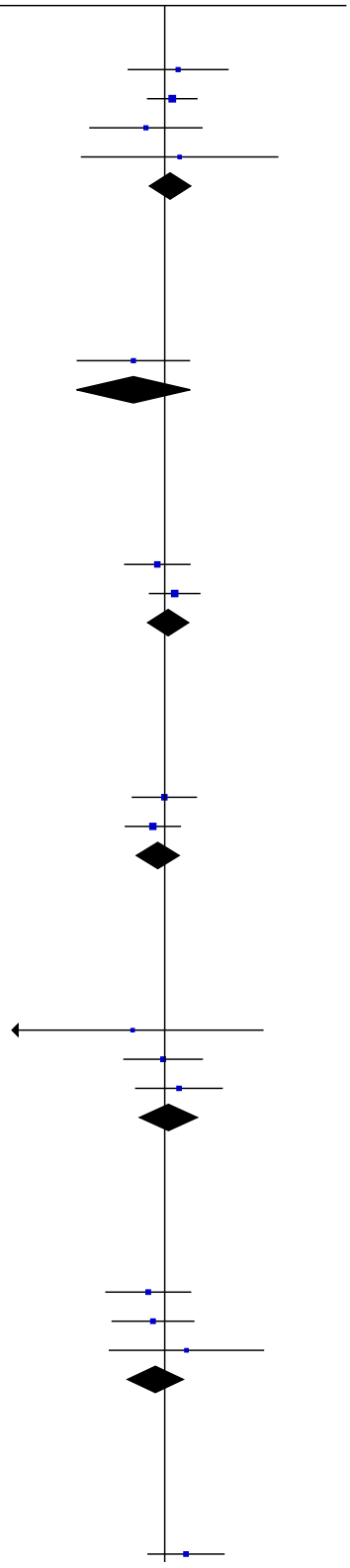
Analysis 7.4. (Continued)

Test for overall effect: Z = 1.53 (P = 0.13)						
7.4.7 vorozole 2.5 mg						
Goss 1999	14	185	20	190	4.5%	0.70 [0.34 , 1.42]
Subtotal (95% CI)		**185**		**190**	**4.5%**	**0.70 [0.34 , 1.42]**
Total events:	14		20			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.99 (P = 0.32)						
Total (95% CI)		**2969**		**2968**	**100.0%**	**0.98 [0.86 , 1.13]**
Total events:	490		497			
Heterogeneity: Chi² = 16.00, df = 17 (P = 0.52); I² = 0%						
Test for overall effect: Z = 0.21 (P = 0.83)						
Test for subgroup differences: Chi² = 6.14, df = 6 (P = 0.41), I² = 2.3%						



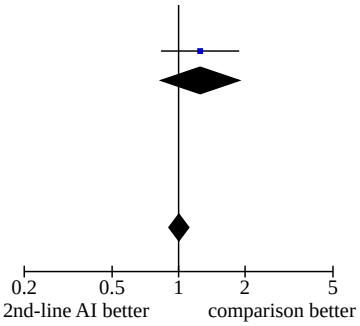
Analysis 7.5. Comparison 7: AI as second-line therapy versus any other therapy, Outcome 5: Clinical benefit (randomised)

Study or Subgroup	comparison Events	Total	2nd line AI Events	Total	Weight	Odds Ratio M-H, Fixed, 95% CI
7.5.1 aminoglutethimide (any dose)						
Canney 1988	61	112	54	106	3.5%	1.15 [0.68 , 1.96]
Garcia-Giralt 1992	231	431	224	434	14.3%	1.08 [0.83 , 1.41]
Lundgren 1989	49	90	51	86	3.3%	0.82 [0.45 , 1.49]
Mercer 1993	13	33	10	28	0.9%	1.17 [0.41 , 3.32]
Subtotal (95% CI)		**666**		**654**	**22.0%**	**1.06 [0.85 , 1.31]**
Total events:	354		339			
Heterogeneity: Chi² = 0.86, df = 3 (P = 0.84); I² = 0%						
Test for overall effect: Z = 0.51 (P = 0.61)						
7.5.2 formestane 250 mg biweekly						
Thuerlimann 1997	46	86	56	91	3.5%	0.72 [0.40 , 1.31]
Subtotal (95% CI)		**86**		**91**	**3.5%**	**0.72 [0.40 , 1.31]**
Total events:	46		56			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.08 (P = 0.28)						
7.5.3 anastrozole 1 mg						
Buzdar 1996a	102	253	111	263	9.0%	0.93 [0.65 , 1.31]
Mauriac 2003	186	428	173	423	13.6%	1.11 [0.85 , 1.46]
Subtotal (95% CI)		**681**		**686**	**22.6%**	**1.04 [0.84 , 1.29]**
Total events:	288		284			
Heterogeneity: Chi² = 0.65, df = 1 (P = 0.42); I² = 0%						
Test for overall effect: Z = 0.33 (P = 0.74)						
7.5.4 exemestane 25 mg						
Chia 2008	87	351	85	342	9.0%	1.00 [0.71 , 1.41]
Kaufmann 2000	135	403	133	366	12.8%	0.88 [0.66 , 1.19]
Subtotal (95% CI)		**754**		**708**	**21.8%**	**0.93 [0.74 , 1.16]**
Total events:	222		218			
Heterogeneity: Chi² = 0.27, df = 1 (P = 0.60); I² = 0%						
Test for overall effect: Z = 0.64 (P = 0.52)						
7.5.5 fadrozole 2 mg						
Bezwoda 1998	4	50	5	46	0.7%	0.71 [0.18 , 2.84]
Buzdar 1996b	65	184	70	196	6.1%	0.98 [0.65 , 1.50]
Buzdar 1996c	61	151	56	152	4.6%	1.16 [0.73 , 1.85]
Subtotal (95% CI)		**385**		**394**	**11.3%**	**1.04 [0.77 , 1.41]**
Total events:	130		131			
Heterogeneity: Chi² = 0.58, df = 2 (P = 0.75); I² = 0%						
Test for overall effect: Z = 0.25 (P = 0.80)						
7.5.6 letrozole 2.5 mg						
Buzdar 2001	47	201	53	199	5.6%	0.84 [0.53 , 1.32]
Dombernowsky 1998	60	189	60	174	5.9%	0.88 [0.57 , 1.37]
Schmid 2001	19	60	14	52	1.4%	1.26 [0.55 , 2.85]
Subtotal (95% CI)		**450**		**425**	**13.0%**	**0.91 [0.68 , 1.21]**
Total events:	126		127			
Heterogeneity: Chi² = 0.73, df = 2 (P = 0.69); I² = 0%						
Test for overall effect: Z = 0.66 (P = 0.51)						
7.5.7 vorozole 2.5mg						
Goss 1999	71	227	60	225	5.7%	1.25 [0.83 , 1.88]



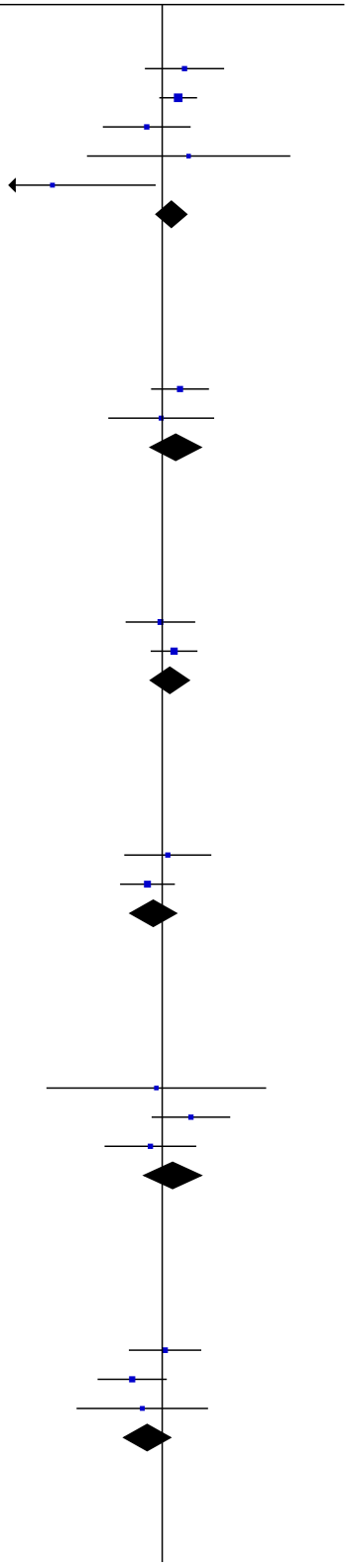
Analysis 7.5. (Continued)

7.5.7 vorozole 2.5mg						
Goss 1999	71	227	60	225	5.7%	1.25 [0.83 , 1.88]
Subtotal (95% CI)		**227**		**225**	**5.7%**	**1.25 [0.83 , 1.88]**
Total events:	71		60			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.08 (P = 0.28)						
Total (95% CI)		**3249**		**3183**	**100.0%**	**1.00 [0.90 , 1.11]**
Total events:	1237		1215			
Heterogeneity: Chi² = 6.71, df = 15 (P = 0.97); I² = 0%						
Test for overall effect: Z = 0.05 (P = 0.96)						
Test for subgroup differences: Chi² = 3.61, df = 6 (P = 0.73), I² = 0%						



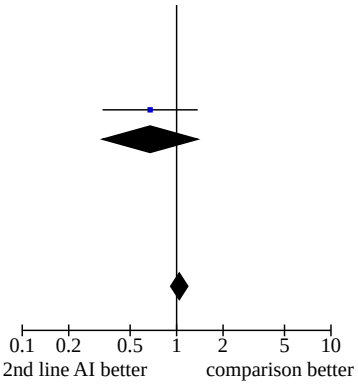
Analysis 7.6. Comparison 7: AI as second-line therapy versus any other therapy, Outcome 6: Objective response (randomised)

Study or Subgroup	comparison Events	Total	2nd line AI Events	Total	Weight	Odds Ratio M-H, Fixed, 95% CI
7.6.1 aminoglutethimide (any dose)						
Canney 1988	35	112	26	106	4.0%	1.40 [0.77 , 2.54]
Garcia-Giralt 1992	154	431	132	434	18.3%	1.27 [0.96 , 1.69]
Lundgren 1989	23	90	26	86	4.3%	0.79 [0.41 , 1.53]
Mercer 1993	5	33	3	28	0.6%	1.49 [0.32 , 6.87]
Russell 1997	2	75	10	80	2.0%	0.19 [0.04 , 0.91]
Subtotal (95% CI)		**741**		**734**	**29.1%**	**1.15 [0.91 , 1.45]**
Total events:	219		197			
Heterogeneity: Chi² = 7.34, df = 4 (P = 0.12); I² = 46%						
Test for overall effect: Z = 1.17 (P = 0.24)						
7.6.2 formestane 250 mg biweekly						
Freue 2000	55	271	45	276	7.7%	1.31 [0.85 , 2.02]
Thuerlimann 1997	14	86	15	91	2.6%	0.99 [0.44 , 2.18]
Subtotal (95% CI)		**357**		**367**	**10.3%**	**1.22 [0.84 , 1.79]**
Total events:	69		60			
Heterogeneity: Chi² = 0.37, df = 1 (P = 0.54); I² = 0%						
Test for overall effect: Z = 1.04 (P = 0.30)						
7.6.3 anastrozole 1 mg						
Buzdar 1996a	31	253	33	263	6.1%	0.97 [0.58 , 1.64]
Mauriac 2003	82	428	70	423	12.3%	1.20 [0.84 , 1.70]
Subtotal (95% CI)		**681**		**686**	**18.4%**	**1.12 [0.84 , 1.50]**
Total events:	113		103			
Heterogeneity: Chi² = 0.41, df = 1 (P = 0.52); I² = 0%						
Test for overall effect: Z = 0.77 (P = 0.44)						
7.6.4 exemestane 25 mg						
Chia 2008	20	351	18	342	3.7%	1.09 [0.56 , 2.09]
Kaufmann 2000	50	403	55	366	10.9%	0.80 [0.53 , 1.21]
Subtotal (95% CI)		**754**		**708**	**14.6%**	**0.87 [0.62 , 1.24]**
Total events:	70		73			
Heterogeneity: Chi² = 0.60, df = 1 (P = 0.44); I² = 0%						
Test for overall effect: Z = 0.76 (P = 0.45)						
7.6.5 fadrozole 2 mg						
Bezwoda 1998	3	50	3	46	0.6%	0.91 [0.18 , 4.78]
Buzdar 1996b	30	184	22	196	3.9%	1.54 [0.85 , 2.78]
Buzdar 1996c	17	151	20	152	3.8%	0.84 [0.42 , 1.67]
Subtotal (95% CI)		**385**		**394**	**8.3%**	**1.17 [0.76 , 1.80]**
Total events:	50		45			
Heterogeneity: Chi² = 1.82, df = 2 (P = 0.40); I² = 0%						
Test for overall effect: Z = 0.71 (P = 0.48)						
7.6.6 letrozole 2.5 mg						
Buzdar 2001	30	180	32	199	5.5%	1.04 [0.61 , 1.80]
Dombernowsky 1998	31	189	41	174	7.7%	0.64 [0.38 , 1.07]
Schmid 2001	9	60	10	52	2.0%	0.74 [0.28 , 1.99]
Subtotal (95% CI)		**429**		**425**	**15.1%**	**0.80 [0.56 , 1.13]**
Total events:	70		83			
Heterogeneity: Chi² = 1.68, df = 2 (P = 0.43); I² = 0%						
Test for overall effect: Z = 1.27 (P = 0.21)						



Analysis 7.6. (Continued)

Test for overall effect: Z = 1.27 (P = 0.21)						
7.6.7 vorozole 2.5 mg						
Goss 1999	14	227	20	225	4.1%	0.67 [0.33 , 1.37]
Subtotal (95% CI)		**227**		**225**	**4.1%**	**0.67 [0.33 , 1.37]**
Total events:	14		20			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.09 (P = 0.28)						
Total (95% CI)		**3574**		**3539**	**100.0%**	**1.04 [0.92 , 1.18]**
Total events:	605		581			
Heterogeneity: Chi² = 19.00, df = 17 (P = 0.33); I² = 11%						
Test for overall effect: Z = 0.61 (P = 0.54)						
Test for subgroup differences: Chi² = 6.56, df = 6 (P = 0.36), I² = 8.5%						



ADDITIONAL TABLES

Table 1. Aromatase inhibitors - description

Generic Name	Trade Name	Generation	Doses used
aminoglutethimide		First	125 mg, 250 mg, 500 mg, 750 mg, 1000 mg
anastrozole	Arimidex	Third, non-steroidal	1 mg, 10 mg
atamestane		Third, steroidal	500mg
exemestane	Aromasin	Third, steroidal	25 mg
fadrozole	CGS16949A	Third, non-steroidal	2 mg
formestane	Lentaron	Second	250 mg im
letrozole	Femara	Third, non-steroidal	0.5 mg, 2 mg, 2.5 mg, 10 mg
vorozole		Third, non-steroidal	2.5 mg

APPENDICES

Appendix 1. The Cochrane Central Register of Controlled Trials (CENTRAL) Issue 2, 2008

Search terms "aromatase inhibitor" AND "randomised trial" AND "breast cancer " AND (advanced OR metastatic)

WHAT'S NEW

Date	Event	Description
7 February 2018	Review declared as stable	Although many trials have been published since the last review version, the review contains sufficient evidence on the efficacy

Date	Event	Description
		of Aromatase inhibitors (AIs) such that AIs are standard therapy. Therefore we do not expect to update this review.

HISTORY

Protocol first published: Issue 4, 2001
Review first published: Issue 1, 2007

Date	Event	Description
6 August 2009	New search has been performed	New search and addition of 7 trials
6 August 2009	New citation required but conclusions have not changed	Update for Issue 4, 2009
5 August 2009	Amended	reference added
8 July 2009	Amended	edited to address additional reviewers' comments
29 June 2009	Amended	edited to address reviewers' comments
3 February 2009	Amended	Feedback from group incorporated
27 November 2008	New search has been performed	Search run by BCG on 2 November 2007. Authors updated search to 30 June 2008. Additional studies identified and data updated
5 September 2008	Amended	Converted to new review format.
13 August 2008	New citation required and conclusions have changed	Substantive amendment

CONTRIBUTIONS OF AUTHORS

Judith Bliss wrote the original protocol, initiated the review with Lorna Gibson and reviewed the review.

Lorna Gibson worked with the Cochrane Breast Group to identify the initial list of references; she then worked through the list to identify eligible trials. Lorna carried out the first independent data extraction from the eligible trials and was a main contributor to the review and analysis, and the update.

Claire Dawson carried out a second independent extraction of data from the eligible trials and contributed to the original review.

David Lawrence carried out an independent check of the data extraction for accuracy and consistency and was a main contributor to the original review analysis, and the update.

DECLARATIONS OF INTEREST

One of the authors (JMB) is a member of the management group and grant holder for the Intergroup Exemestane Study. This is funded by Pfizer, the producers of the aromatase inhibitor exemestane.

SOURCES OF SUPPORT

Internal sources

- Cancer Research UK, UK

External sources

- No sources of support supplied

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

Many of the data required to carry out analyses of prospectively identified subgroups, as set out in the review protocol, were not available. We could not, therefore, identify specific subgroups of women who may benefit from AI use.

NOTES

This updated review includes the following additional seven trials to the 30 in the original publication: Chia 2008; Gale 1994; Garcia-Giralt 1992; Goss 2007; Lundgren 1989; Samonis 1994. There were also two papers by Mourisden and colleagues (Mourisden 2004; Mourisden 2007) which contributed follow-up information.

The update demonstrated a survival benefit of 10% with the use of AIs for the treatment of advanced (metastatic) breast cancer, compared to 11% in the original review.

INDEX TERMS

Medical Subject Headings (MeSH)

Antineoplastic Agents, Hormonal [*therapeutic use]; Aromatase Inhibitors [*therapeutic use]; Breast Neoplasms [*drug therapy] [mortality]; Neoplasms, Hormone-Dependent [*drug therapy]; *Postmenopause; Randomized Controlled Trials as Topic

MeSH check words

Female; Humans