

RESEARCH ARTICLE

Open Access

# A process for assessing the feasibility of a network meta-analysis: a case study of everolimus in combination with hormonal therapy versus chemotherapy for advanced breast cancer

Shannon Cope<sup>1\*</sup>, Jie Zhang<sup>2</sup>, Stephen Saletan<sup>2</sup>, Brielan Smiechowski<sup>3</sup>, Jeroen P Jansen<sup>4</sup> and Peter Schmid<sup>5</sup>

## Abstract

**Background:** The aim of this study is to outline a general process for assessing the feasibility of performing a valid network meta-analysis (NMA) of randomized controlled trials (RCTs) to synthesize direct and indirect evidence for alternative treatments for a specific disease population.

**Methods:** Several steps to assess the feasibility of an NMA are proposed based on existing recommendations. Next, a case study is used to illustrate this NMA feasibility assessment process in order to compare everolimus in combination with hormonal therapy to alternative chemotherapies in terms of progression-free survival for women with advanced breast cancer.

**Results:** A general process for assessing the feasibility of an NMA is outlined that incorporates explicit steps to visualize the heterogeneity in terms of treatment and outcome characteristics (Part A) as well as the study and patient characteristics (Part B). Additionally, steps are performed to illustrate differences within and across different types of direct comparisons in terms of baseline risk (Part C) and observed treatment effects (Part D) since there is a risk that the treatment effect modifiers identified may not explain the observed heterogeneity or inconsistency in the results due to unexpected, unreported or unmeasured differences. Depending on the data available, alternative approaches are suggested: list assumptions, perform a meta-regression analysis, subgroup analysis, sensitivity analyses, or summarize why an NMA is not feasible.

**Conclusions:** The process outlined to assess the feasibility of an NMA provides a stepwise framework that will help to ensure that the underlying assumptions are systematically explored and that the risks (and benefits) of pooling and indirectly comparing treatment effects from RCTs for a particular research question are transparent.

**Keywords:** Advanced breast cancer, Everolimus, Chemotherapy, Network meta-analysis, Progression-free survival, Systematic literature review, Feasibility assessment

## Background

Network meta-analyses (NMA) are increasingly being performed to inform decision-making regarding the comparative efficacy and safety of alternative treatments [1]. In order to determine the comparative efficacy or safety of a new treatment using a NMA it is necessary to establish the relevant comparators. Generally, the indication for the

new treatment and the way in which the new treatment is expected to be used in clinical practice will determine the comparators of interest. In some cases the comparators are explicitly defined by reimbursement agencies for a technology appraisal, which is the case in the United Kingdom where the National Institute for Health and Care Excellence (NICE) develops a final scope based on a stakeholder consultation process [2].

In order to inform decision-making it is necessary to assess whether it is feasible to perform a valid NMA to compare the new treatment with usual care based on

\* Correspondence: scope@mapigroup.com

<sup>1</sup>Mapi, 33 Bloor Street East, Suite 1300, Toronto, Ontario M4W 3H1, Canada  
Full list of author information is available at the end of the article

the available randomized controlled trials (RCTs). As with any NMA, the validity of such analysis relies on whether there are systematic differences among the studies included in the network across treatment comparisons, especially patient or disease characteristics that are treatment effect modifiers [3-6]. Although there is guidance available regarding the underlying principles of an NMA, there is a need for a more structured process that incorporates both clinical and methodological expertise to assess the feasibility of performing a valid NMA [7]. The aim of this study is to outline a general process for assessing the feasibility of performing a valid NMA. A case study is used to illustrate the feasibility of performing an NMA to compare everolimus in combination with hormonal therapy to alternative chemotherapies in terms of progression-free survival (PFS) for women with advanced breast cancer (ABC).

The first section presents general steps for assessing the feasibility of a NMA. Next, the case study is presented in terms of the background, the identification and selection of trials, the method for the systematic review and analysis, and the results of the feasibility assessment and NMA. Readers are encouraged to use our application of these rules to our clinical example as a case study in applying our process to a possible research question and may envision ways to apply it to their own review.

## Methods

### Assessing feasibility of a network meta-analysis

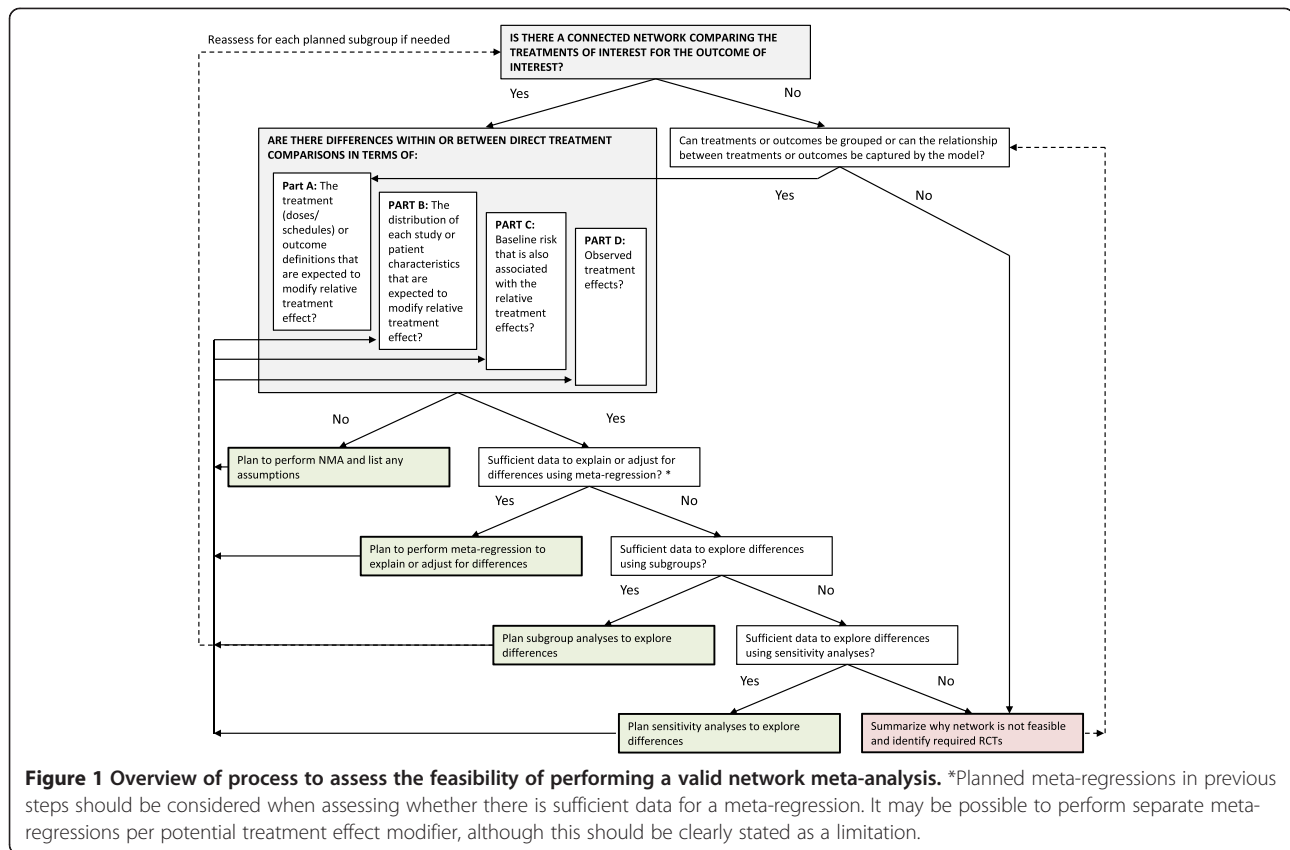
In the absence of trials involving a direct comparison of interventions, an indirect comparison can provide valuable evidence for the relative treatment effects between competing interventions [5,8-14]. Even when the results of the direct evidence are conclusive, combining them with the results of indirect estimates in a mixed treatment comparison may yield more precise estimates as a greater evidence base is considered [8,10,12]. If the available evidence base consists of a network of interlinked multiple RCTs involving treatments compared directly or indirectly or both, it can be synthesized by means of NMA [4,8,11,13,14]. Since randomization of patients does not hold across trials in a network of RCTs, there might be differences across treatments that may compromise the validity of a NMA. It can be expected that there will always be some degree of variation in patient characteristics across studies. If these characteristics are effect modifiers of the relative treatment effects of interest then there will be heterogeneity in the evidence base. If there is an imbalance in relative treatment effect modifiers across comparisons, then the transitivity and consistency assumptions do not hold and some or all of the estimates of the NMA will be biased [5,9,11]. Therefore, it is important to assess whether there are differences

in study and patient characteristics across comparisons that affect the summary measures of treatment effects (that is, odds ratio or hazard ratio) for the interventions of interest relative to an overall reference treatment. (Note: Differences in prognostic factors that are not also treatment effect modifiers do not impact the validity of the analysis).

A general process for assessing the feasibility of an NMA is outlined in Figure 1, which builds upon the existing recommendations regarding NMA [5,6,11]. Initially, steps to visualize the clinical heterogeneity in terms of treatment and outcome characteristics (Part A) as well as the study and patient characteristics (Part B) are proposed [7]. Next, steps are suggested to assess differences within and across the direct pairwise comparisons in terms of baseline risk (Part C) and observed treatment effects (Part D) since there is a risk that the treatment effect modifiers identified may not explain the observed heterogeneity or inconsistency in the results due to unexpected, unreported or unmeasured differences.

The proposed process for the feasibility assessment of an NMA is recommended to be specified in the systematic review protocol and statistical analysis plan. It is recommended to develop a parsimonious list of potential treatment effect modifiers for the study and patient characteristics based on prior knowledge before beginning the systematic review and feasibility assessment that will help guide Part B [6]. The process also acknowledges that some additional differences may be identified that act as treatment effect modifiers, which could be predefined in terms of a process to identify outliers. In some cases a modification to the statistical analysis plan may be required. For example, decisions regarding pooling different treatment doses or regimens (Part A) may be challenging to pre-specify in the statistical analysis plan without having identified the relevant evidence.

It is recommended that the decision to perform an NMA should be based primarily on clinical judgment of whether differences among studies may affect the comparisons of treatments or make some comparisons inappropriate [11]. However, an evaluation of baseline risk and heterogeneity (or inconsistency) in observed treatment effects in Part C and Part D may help identify analyses to adjust for differences [4,15]. Therefore, it is suggested to pre-specify the types of analyses that will be used to explore heterogeneity and/or inconsistency, which may include the use of a random effects model, unrelated means model [4], exclusion of specific studies that are outliers (using pre-defined criteria), node splitting [16], or the inclusion of inconsistency factors [17-19]. Ultimately, using this process will help to ensure that the risks (and benefits) of pooling and indirectly comparing treatment effects reported in RCTs for a particular research question are clearly documented [11].



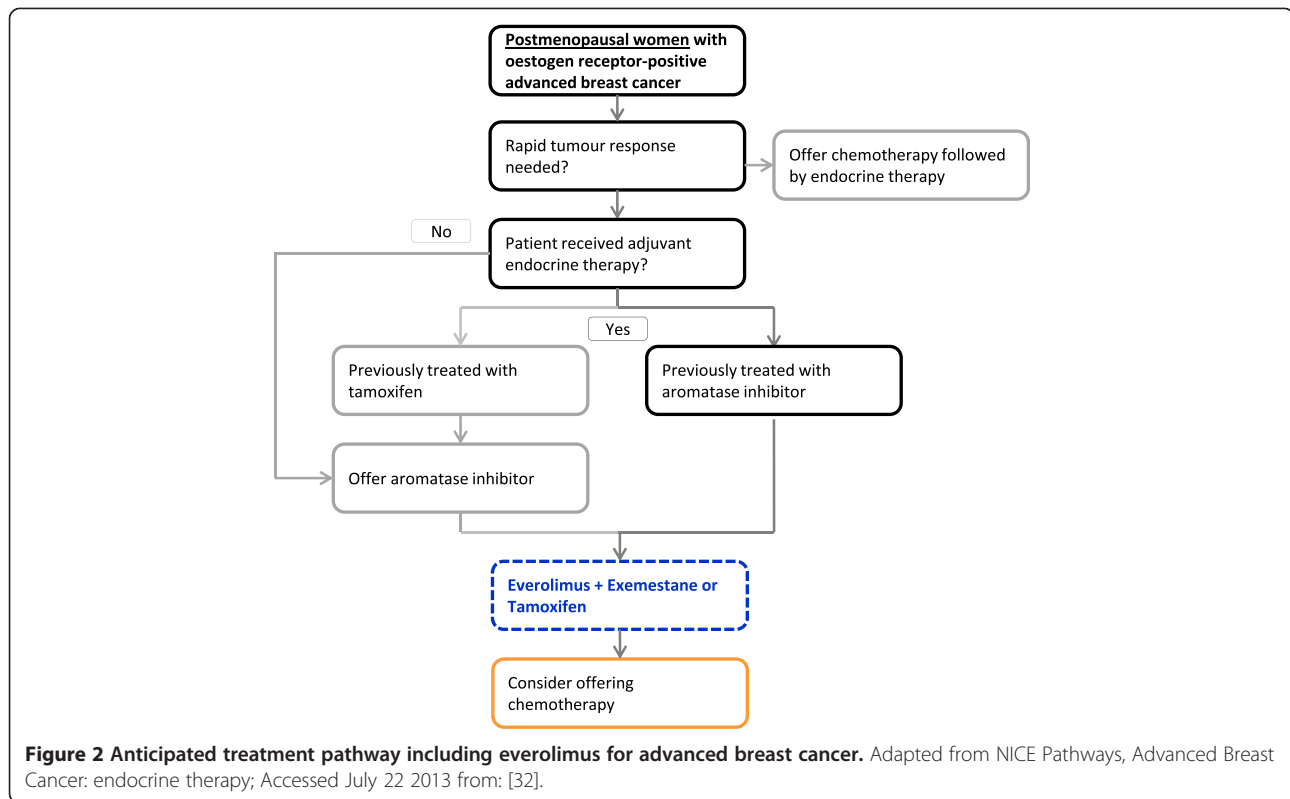
An important step in the feasibility assessment involves making a judgment on whether or not there is sufficient data to explain or adjust for differences in potential treatment effect modifiers using a meta-regression. Although there is no clear threshold regarding the number of data points required to perform a meta-regression, Gagnier *et al.* refer to a general rule of thumb which suggests that there should be close to ten trials when working with summary or aggregate patient data (or ten individuals per variable, when working with pooled or individual patient data), but also caution that fewer studies may be associated with more heterogeneity, and more variables explored may be associated with a higher type 1 error rate [7]. Therefore, the number of data points should be considered against the number of parameters included in the meta-regression model. Generally, when these types of meta-regressions are based on aggregate study level data, it is assumed that the effect of the covariate is constant across the different treatments, which is recommended for most cases given the limited data available [20]. While accounting for study-level factors using these types of models is advised, the risk of ecological bias should be recognized when adjusting for differences in patient characteristics in the absence of individual patient level data [21,22].

### Case study

Everolimus is indicated for the treatment of hormone receptor-positive, human epidermal growth factor receptor type 2 negative (HER2-) ABC, in combination with exemestane (EXE), in postmenopausal women without symptomatic visceral disease after recurrence or progression following a non-steroidal aromatase inhibitor (NSAIs) [23].

The phase III RCT BOLERO-2 demonstrated that everolimus plus EXE more than doubled median PFS compared with placebo plus EXE while still maintaining quality of life in ABC patients (estrogen receptor positive (ER+) and HER2-) who recurred or progressed during or after NSAIs [24-27]. Additionally, the phase II TAMRAD trial demonstrated the efficacy of everolimus in combination with tamoxifen (TAM) in comparison to TAM alone [28].

For women with ER+ ABC, guidelines recommend endocrine therapy as the preferred option even in the presence of visceral disease. If there is evidence of endocrine resistance or rapidly progressive disease requiring a fast response then chemotherapy is recommended [29]. Figure 2 illustrates the current treatment pathway for ER+ patients with ABC and also outlines the anticipated use of everolimus in the treatment sequence. Everolimus



in combination with EXE offers a viable new line of therapy which can delay treatment with chemotherapy and has been described as ‘a step change in treatment’ [30]. Consequently the comparison of everolimus in combination with hormonal therapy versus chemotherapy is of interest and was defined as a comparator of interest by NICE [31].

#### Identification and selection of studies

A systematic literature search was performed in March 2013 to identify published RCTs evaluating the efficacy of treatment regimens for patients with postmenopausal ABC (stage III or IV) who were treated with everolimus, alternative hormonal therapies or alternative chemotherapies to facilitate an indirect comparison of everolimus versus chemotherapy. Medline, Medline In-Process, EMBASE and Cochrane databases were searched by using a predefined search strategy with terms relevant to ABC, RCTs and the comparisons of interest (see search strategy in Additional file 1).

Two reviewers independently evaluated each identified study against the following predetermined criteria:

Population: postmenopausal women with ABC (locally ABC (stage III) or metastatic breast cancer (stage IV)).  
 Interventions: everolimus (in combination with hormonal therapy) and chemotherapies including capecitabine, vinorelbine, paclitaxel, docetaxel, nab-

paclitaxel, doxorubicin, epirubicin, pegylated liposomal doxorubicin and eribulin.

Comparisons Step 1: everolimus (in combination with hormonal therapy) versus hormonal therapy alone; EXE versus TAM; hormonal therapy versus chemotherapy.

Comparisons Step 2: chemotherapy identified in Step 1 versus any chemotherapies of interest; Comparison of alternative hormonal therapies identified in Step 1.

Outcomes: Kaplan Meier curves reporting PFS, time to progression (TTP) or time to failure (TTF).

Study design: phase II or III RCTs.

Although the target population for everolimus was based on the BOLERO-2 trial, which included only ER+ HER2- patients with prior NSAI treatments, the scope of the population was defined more broadly in order to capture all available evidence for trials comparing hormonal therapy to chemotherapy.

For each identified study that met the selection criteria, details were extracted on study design, study population characteristics and interventions. For all studies, the reported PFS Kaplan-Meier curves were digitized for each treatment arm (DigitizeIt v1.6.1). The data set was created on the basis of extracted progression proportions (including PFS, TTP or TTF), which were used to calculate the incident number of events for each interval and patients at risk at the beginning of that interval [33].

Generally, PFS is defined as the time elapsed between randomization and tumor progression or death from any cause, with censoring of patients who are lost to follow-up, where progression events include an increase in tumor size and/or the development of new tumors according to standardized criteria such as Response Evaluation Criteria in Solid Tumors. TTP is typically defined as the time elapsed between randomization and tumor progression, with censoring of patients who die or are lost to follow-up. Finally, TTF may be defined as the time from randomization until progression, relapse, or death from any cause, although this may vary. PFS, TTP and TTF are not always labelled or defined consistently in different trials [34].

#### **Assessing the feasibility of a network meta-analysis for postmenopausal women with advanced breast cancer**

The process outlined in Figure 1 was applied in order to assess whether a NMA was feasible to indirectly compare everolimus (in combination with hormonal therapy) to alternative chemotherapies in terms of PFS for postmenopausal women with ABC. The following potential treatment effect modifiers were identified *a priori* based on clinical expertise: hormone receptor status (HR-status), prior hormonal therapy, prior chemotherapy, visceral metastases, performance status and age. For parts A and B, network diagrams illustrating the structure of the network as well as differences in outcome definitions and potential treatment effect modifiers were developed. For parts C and D the baseline risk and heterogeneity in observed treatment effects were also illustrated to facilitate an assessment of the differences within and across direct treatment comparisons.

A Bayesian NMA was planned using the methodology introduced by Ouwens *et al.* and Jansen *et al.* to synthesize and indirectly compare the published PFS Kaplan-Meier curves for each treatment and RCT [33,35,36]. With this approach, the PFS of patients over time of the interventions compared in a trial is modeled with parametric survival functions and the difference in the shape and scale parameters of these functions between interventions is synthesized and indirectly compared across trials. The best fitting first order fractional polynomial model was selected [36] using the deviance information criteria (DIC) [37,38]. Additional details on these models and developing the datasets have been reported previously [33,35,36,39].

## **Results**

### **Study identification**

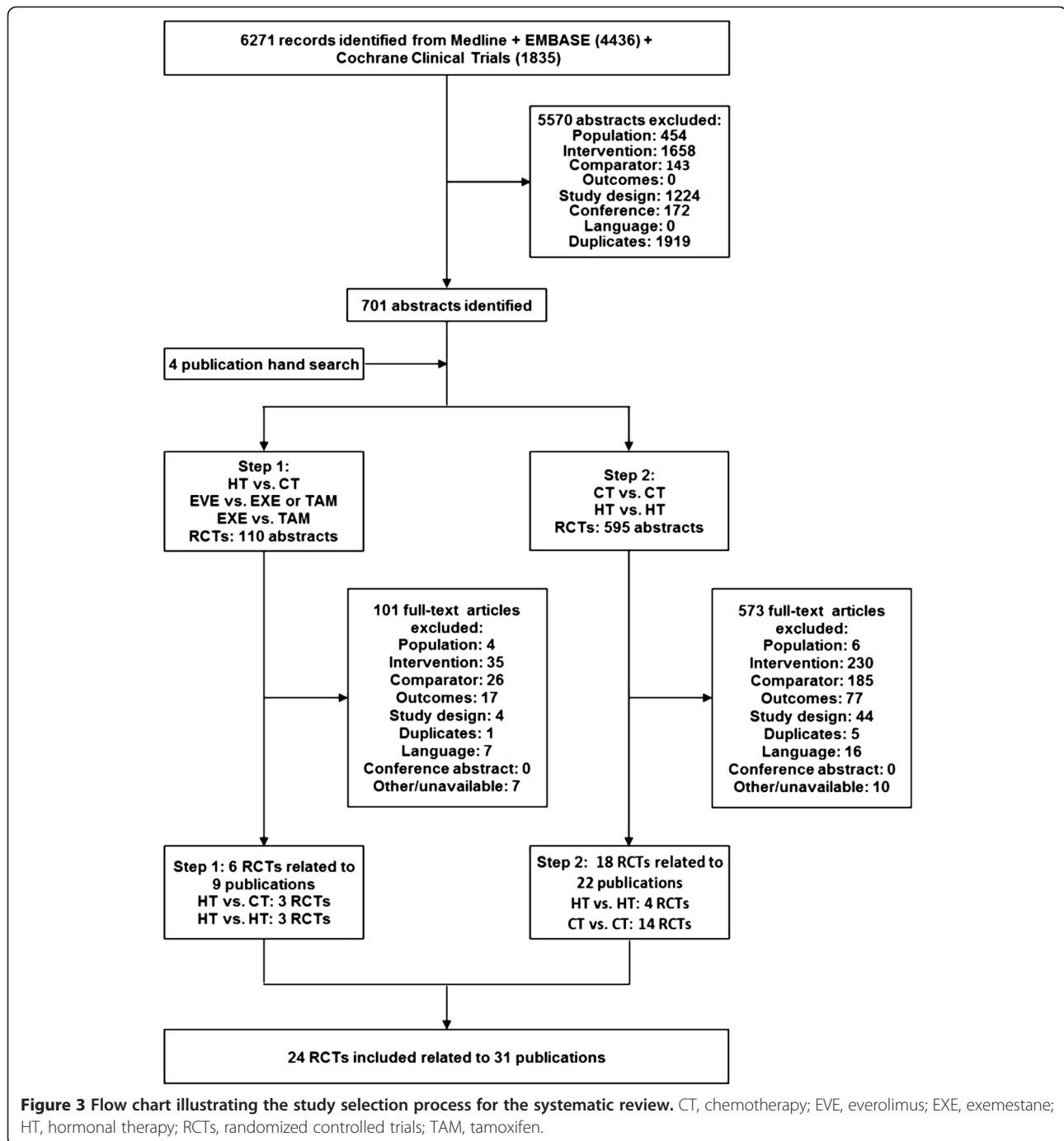
A stepwise process was used to select the relevant RCTs and to assess the feasibility of a NMA (Figure 3 presents the flow chart, while Additional file 2: Table S1 presents all included trials, their respective authors, years of publication and interventions) [24-28,40-64]. Based on

the first step of the systematic literature review three RCTs were identified comparing hormonal therapies to chemotherapies (TAM versus cyclophosphamide + doxorubicin (CD), megestrol acetate (MA) versus cyclophosphamide + methotrexate + fluorouracil (CMF), and MA versus mitoxantrone (MZ) [40,43,64]). These RCTs were connected to three RCTs identified comparing everolimus + TAM to TAM, everolimus + EXE to EXE, and TAM to EXE [24,25,28,59-62]. Therefore, it was possible to develop a network of connected RCTs based on these six RCTs. However, this network, based on step 1, only allowed for an indirect comparison of everolimus (plus hormonal therapy) to the chemotherapies CD and CMF. Since these chemotherapy combinations were not of interest, the network was extended to include an additional 14 RCTs comparing alternative chemotherapies of interest [41,42,45-47,49,50,54,57-60,63,65,66].

Additionally, four RCTs comparing alternative hormonal therapies in the network were included to strengthen the network (comparing MA to TAM or EXE) [43,48,51-53,55,56] based on the final step outlined in the systematic review protocol. Although the objective of this network was not to compare the efficacy of alternative hormonal therapies, these RCTs provided additional connections between trials evaluating everolimus and those treatments required to link to chemotherapies of interest. As was done for this case study, we would advise the utilization of the most 'comprehensive' evidence base in the initial feasibility assessment. However, for trials comparing treatments that are not directly of interest for the decision-problem it is important to consider that the additional value of evidence should be weighed against the risk of introducing additional heterogeneity (or inconsistency) to the network. Exploring the impact of a broader evidence base for the network was not the purpose of this study, although this is an important issue that has been considered by Cooper *et al.* [67].

### **Is there a connected network comparing the treatments of interest for the outcome of interest?**

The available RCTs formed a connected network of evidence in order to indirectly compare everolimus in combination with hormonal therapy to the relevant chemotherapies, which is illustrated in Additional file 3: Figure S1. However, the indirect comparison of everolimus (in combination with hormonal therapy) versus any of the chemotherapies of interest is mediated by at least four different treatment comparisons. The studies included between 48 to 769 patients per study, with five studies including fewer than 100 patients. Based on the number of relatively small RCTs available and the large 'degree of separation' in the network pathway connecting the treatments of interest, the estimates of the indirect treatment were expected to be very uncertain.



Are there differences within or between direct treatment comparisons for the outcome of interest?

*Part A - The treatment (doses/ schedules) or outcome definitions that are expected to modify relative treatment effect*

Some differences in the treatment doses and/or schedules were identified, which are summarized in Additional file 4: Table S2. Since there were a limited number of RCTs included per treatment (ranging from one up to

seven RCTs for doxorubicin), a meta-regression was not deemed to be feasible to adjust for differences in treatment doses. However, the RCTs that were most different in terms of treatment dose were planned to be excluded in a sensitivity analysis based on clinical expertise. In the 'All Evidence' base case analysis the different treatment doses were grouped together assuming that there would have been no differences in outcomes (beyond sampling error) between the different treatments within

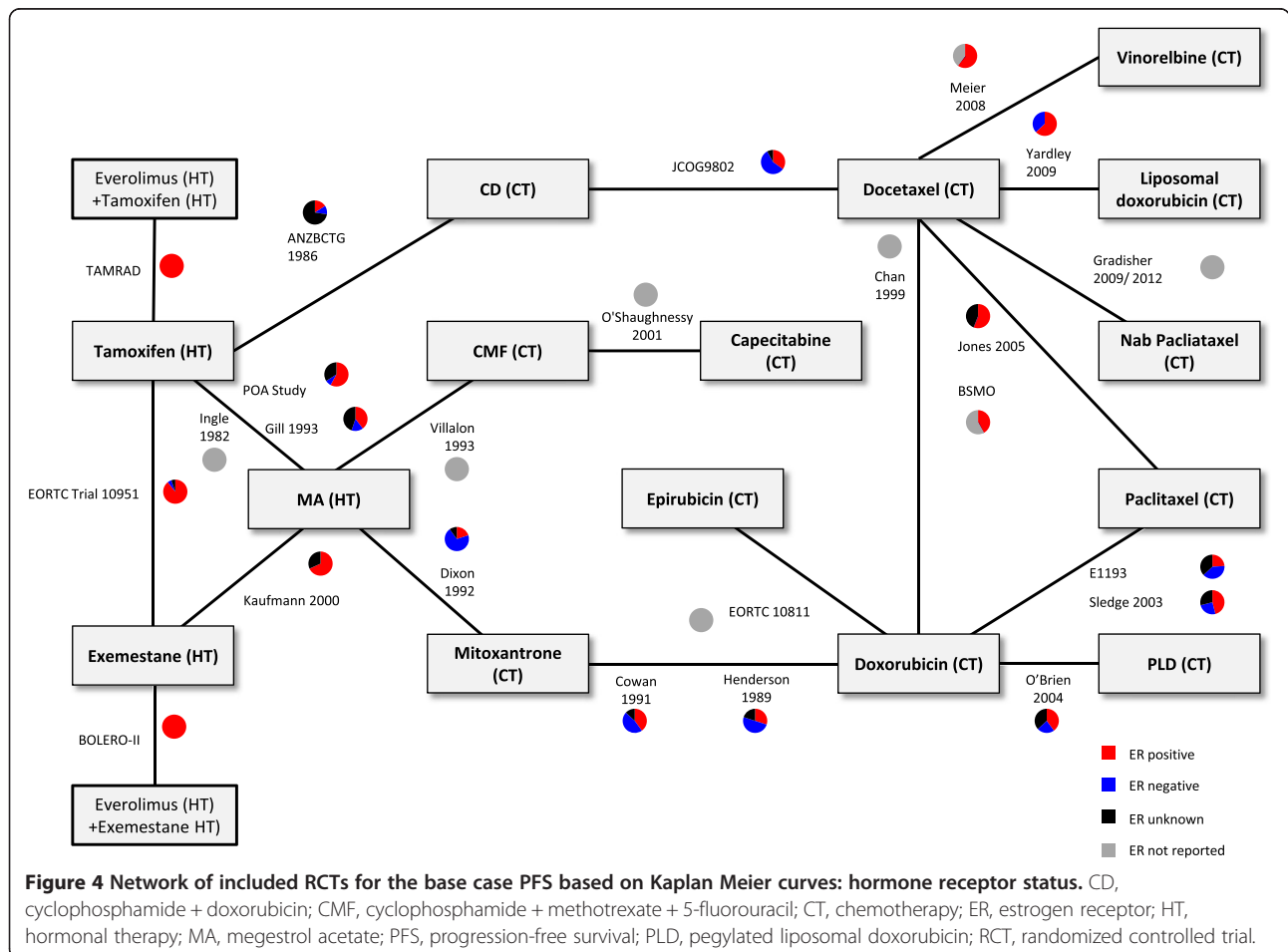
the group if the same population would have been treated. A comparison between the base case analysis and the sensitivity analysis excluding specific treatment doses may be possible to explore the impact of this assumption.

Differences across the RCTs and treatment comparisons were also identified in terms of the outcome definition. Eleven RCTs reported TTP, seven RCTs reported PFS and six RCTs reported TTF (Figure 4). The network included 24 trials evaluating two treatment arms of interest in terms of 17 different treatments. Therefore, the simplest arm-based (fixed effect) model assuming a constant hazard ratio to synthesize the network of studies would include 40 parameters (24 baseline effects for each study and (17-1 = 16) treatment effects), whereas an analysis based on reported hazard ratios would provide only 24 data points. A meta-regression to adjust for differences in the PFS definitions was planned but may not be feasible given the limited amount of data; therefore, a sensitivity analysis was planned to exclude RCTs within the network evaluating TTF [40,44,48,50,63,68], which may differ the most from TTP and PFS.

**Part B – The distribution of study or patient characteristics that are expected to modify relative treatment effects (defined a priori)**

RCTs in the network were generally multicenter open-label trials evaluating women with either advanced or recurrent breast cancer and the study design was considered to be broadly comparable despite some single center trials [43,64]. A summary of the risk of bias across the trials is presented in Additional file 5: Figure S2 and an overview of the risk of bias per RCTs is presented in Additional file 6: Figure S3. Given the limited variation in the blinding of patients across the RCTs, it was determined that there was insufficient data to assess this assumption in terms of a meta-regression. Similarly, since so many of the RCTs were open-label, a sensitivity analysis to exclude these studies was deemed not to be feasible. Therefore, in order to synthesize the trials in the ‘All Evidence’ base case network, it would be necessary to assume that differences in the study design or quality did not act as a treatment effect modifier, which may bias the treatment estimates.

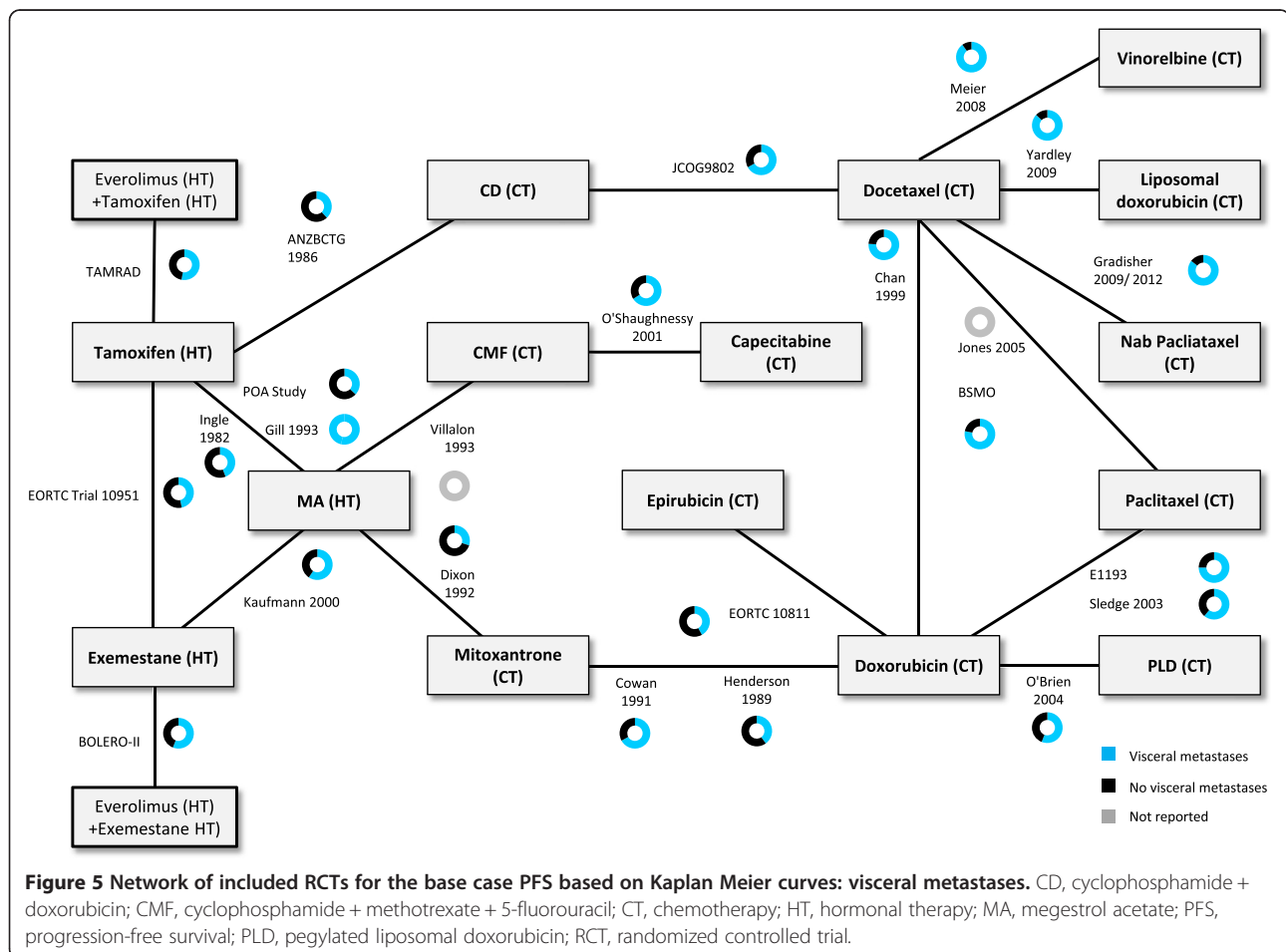
Differences were identified in terms of the distribution of patient characteristics within and across the treatment



comparisons with respect to the potential treatment effect modifiers for ER status (Figure 4), exposure to prior hormonal therapies [see Additional file 7: Figure S4], exposure to prior chemotherapies [see Additional file 8: Figure S5] and visceral metastases (Figure 5). Although a meta-regression to adjust for each of these differences in a separate analysis may be possible, since there were such a large number of studies where the distribution of patient characteristics was not reported, it is likely that these adjustments will be inadequate. Sensitivity analyses to exclude specific studies were not planned given the substantial variation in the differences across the treatment comparisons and the reliance of the network on several older studies comparing hormonal therapy to chemotherapy that did not report several patient characteristics. Therefore, there is a risk that treatment estimates based on an NMA of the base case network will be biased, as the results would only be valid if it can be assumed that differences in the distribution of patient characteristics do not act as treatment effect modifiers (that is, ER status, exposure to prior hormonal therapies,

exposure to prior chemotherapies and visceral metastases). Differences in terms of performance status and age were considered less prominent and are summarized in Table 1. Given the variation observed, sensitivity analyses to exclude outlier studies were planned.

Beyond the differences identified based on the pre-defined potential treatment effect modifiers, differences were also detected in terms of post-menopausal status, HER2 status, and types of prior hormonal therapies, which are also summarized in Table 1. Sensitivity analyses were planned to exclude studies that were outliers in terms of post-menopausal status [48]. However, no sensitivity analyses were possible in terms of HER2 status, as only the everolimus trials reported this characteristic. Similarly, sensitivity analyses were not planned for differences identified in terms of the type of prior hormonal therapy given the limited number of studies that reported this information: patients in the everolimus trials received prior aromatase inhibitor (letrozole or anastrozole), whereas patients in other studies received an estrogen receptor antagonist (TAM) or this information was not reported.





**Table 1 Overview of variation in potential treatment effect modifiers in the RCTs included in the progression-free survival analysis**

Category	Study	Comparison	N	ER+	ER NA or ER unknown	Failed HT adjuvant/metastatic	Failed CT advanced/metastatic	Visceral meta-stases	ECOG 0 or 1	Median age (years)	Post-meno-pausal	HER2-	Type prior HT
<b>Comparisons</b>	Baselga 2012, BOLERO-II [24]	EXE vs. EVE + EXE	724	100%	0%	100%	NR	56%	96%	62	100%	100%	AI <sup>a</sup>
	Bachelot 2012, [28] TAMRAD	TAM vs. EVE + TAM	111	100%	NR	100%	24%	48%	92%	63	100%	95%	AI <sup>a</sup>
HT vs. HT	Paridaens 2008, [60-62] EORTC-10951	EXE vs. TAM	371	89%	5%	22%	33%	47%	87%	62	100%	NR	TAM
HT vs. HT	Kaufmann 2000 [51-53]	EXE vs. MA	769	68%	32%	100%	17%	59%	NR	65	100%	NR	TAM
HT vs. HT	Muss 1985, [55,56] POASTUDY	TAM vs. MA	136	58%	34%	3%	10%	37%	79%	62	100%	NR	NR
HT vs. HT	Ingle 1982 [48]	TAM vs. MA	55	NR	NR	0%	NR	44%	79%	49	48%	NR	NA
HT vs. HT	Gill 1993 [44]	TAM vs. MA	118	40%	45%	0%	0%	53%	NR	NR	100%	NR	NA
HT vs. CT	ANZBCTG 1986 [40]	TAM vs. CD	226	16%	73%	NR	NR	38%	60%	NR	100%	NR	NR
HT vs. CT	Dixon 1992 [43]	MA vs. MZ	60	20%	10%	100%	0%	30%	100%	61	100%	NR	TAM
HT vs. CT	Villalon 1993 [64]	MA vs. CMF	48	NR	NR	NR	0%	NR	NR	NR	83%	NR	NR
CT vs. CT	Cowan 1991 [68]	MZ vs. DOX	237	40%	13%	55%	59%	67%	73%	NR	83%	NR	NR
CT vs. CT	Henderson 1989 [47]	MZ vs. DOX	325	30%	20%	52%	63%	40%	71%	NR	85%	NR	NR
CT vs. CT	Katsumata 2009, [50] JCOG9802	CD vs. DOC	293	35%	8%	100%	0%	67%	95%	54	NR	NR	NR
CT vs. CT	O'Shaughnessy 2001 [58]	CMF vs. CAP	93	NR	NR	91%	0%	66%	NR	70	100%	NR	49% TAM
CT vs. CT	Chan 1999 [65]	DOX vs. DOC	326	NR	NR	71%	58%	76%	NR	NR	NR	NR	NR
CT vs. CT	Paridaens 2000 [59]	DOX vs. PAC	331	24%	37%	74%	0%	75%	91%	55	NR	NR	NR
CT vs. CT	Sledge 2003, [63] E1193	DOX vs. PAC	453	46%	29%	60%	0%	61%	85%	58	NR	NR	NR
CT vs. CT	Bontenbal 1998, [42] EORTC 10811	DOX vs. EPI	232	NR	NR	36%	98%	42%	73%	56	92%	NR	NR
CT vs. CT	O'Brien 2004 [57]	DOX vs. PLD	509	40%	37%	NR	0%	56%	89%	58	62%	NR	NR
CT vs. CT	Jones 2005 [49]	DOC vs. PAC	449	56%	44%	60%	58%	NR	NR	56	88%	NR	NR
CT vs. CT	Beuselinck 2010, [41] BSMO	DOC vs. PAC	70	42%	NR	NR	81%	78%	84%	NR	100%	NR	NR
CT vs. CT	Gradishar 2009 [45,46]	DOC vs. Nab-PAC	148	NR	NR	NR	0%	91%	97%	NR	81%	NR	NR
CT vs. CT	Yardley 2009 [66]	DOC vs. L-DOX	102	63%	0%	60%	0%	87%	91%	63	NR	NR	NR
CT vs. CT	Meier 2008 [54]	DOC vs. VIN	120	60%	NR	NR	90%	90%	74%	60	NR	NR	NR

<sup>a</sup>AI, aromatase inhibitors included letrozole and anastrozole. CD, cyclophosphamide + doxorubicin; CMF, cyclophosphamide + methotrexate + fluorouracil; CAP, capecitabine; CT, chemotherapy; DOC, docetaxel; DOX, doxorubicin; ECOG, European Co-operative Oncology Group; ER, estrogen receptor; EPI, epirubicin; EVE, everolimus; EXE, exemestane; HER2, human epidermal growth factor receptor type 2; HT, hormonal therapy; L-DOX, liposomal doxorubicin; MA, megestrol acetate; MZ, mitozantrone; PAC, paclitaxel; PLD, pegylated liposomal doxorubicin; TAM, tamoxifen; VIN, vinorelbine.

### **Part C - Differences in baseline risk that are associated with the relative treatment effect**

The RCTs in the network were conducted between 1978 and 2012. Therefore, there may be a risk that differences attributable to changes in clinical practice over time could influence the treatment effect, which may justify adjusting for baseline risk. However, in this network there was no clear placebo or standard of care to provide a common treatment arm in order to assess baseline risk. TAM was selected as the baseline comparator of the analysis given the proximity to everolimus and the larger number of RCTs compared to TAM in the network. Additional file 9: Figure S6 presents the individual study results for PFS as extracted from the Kaplan Meier curves per treatment for the included RCTs. This illustration suggests that there is some variation in the baseline risk for TAM, although it is unclear whether these differences are associated with baseline risk. Since there was no clear reference treatment that was common to all (or most) comparisons in the network, adjusting for differences in baseline risk in a meta-regression was not considered. Furthermore, no clear outliers were identified based on the observed variation in baseline risk. Consequently, an NMA combining the RCTs in the base case is at risk of bias due to differences in clinical practice over time, although the role of baseline risk on the treatment effect is not clear.

### **Part D - Differences in observed treatment effects**

Additional file 10: Figure S7 presents the PFS results by study. In the network of evidence there are only four comparisons that are supported by multiple RCTs. For the three RCTs comparing MA versus TAM [44,48,56] the differences consistently favor TAM although there is some variation in the magnitude of the observed effect. For the comparison of doxorubicin versus MZ, both RCTs suggest that there is very little difference in PFS between the treatments [47,68], with doxorubicin showing a slightly more favorable result. However, in the comparison of paclitaxel versus doxorubicin the study by Paridaens *et al.* [59,60] favors doxorubicin over paclitaxel, whereas the study by Sledge *et al.* [63] suggests the opposite, although differences in both cases are minimal. Similarly, for the comparison of paclitaxel versus docetaxel the study by Jones *et al.* [49] suggests docetaxel is favored over paclitaxel, whereas the study by Beuselink *et al.* [41] found very little difference between the treatments and the curves cross each other more than once. Therefore, within the pairwise comparison some heterogeneity was identified, primarily with respect to the chemotherapy comparisons, although the differences were not substantial. The between-study standard deviation resulting from the analysis was 0.41 (95% credible interval (CrI): 0.22, 0.81), which suggests significant between study heterogeneity was present that was

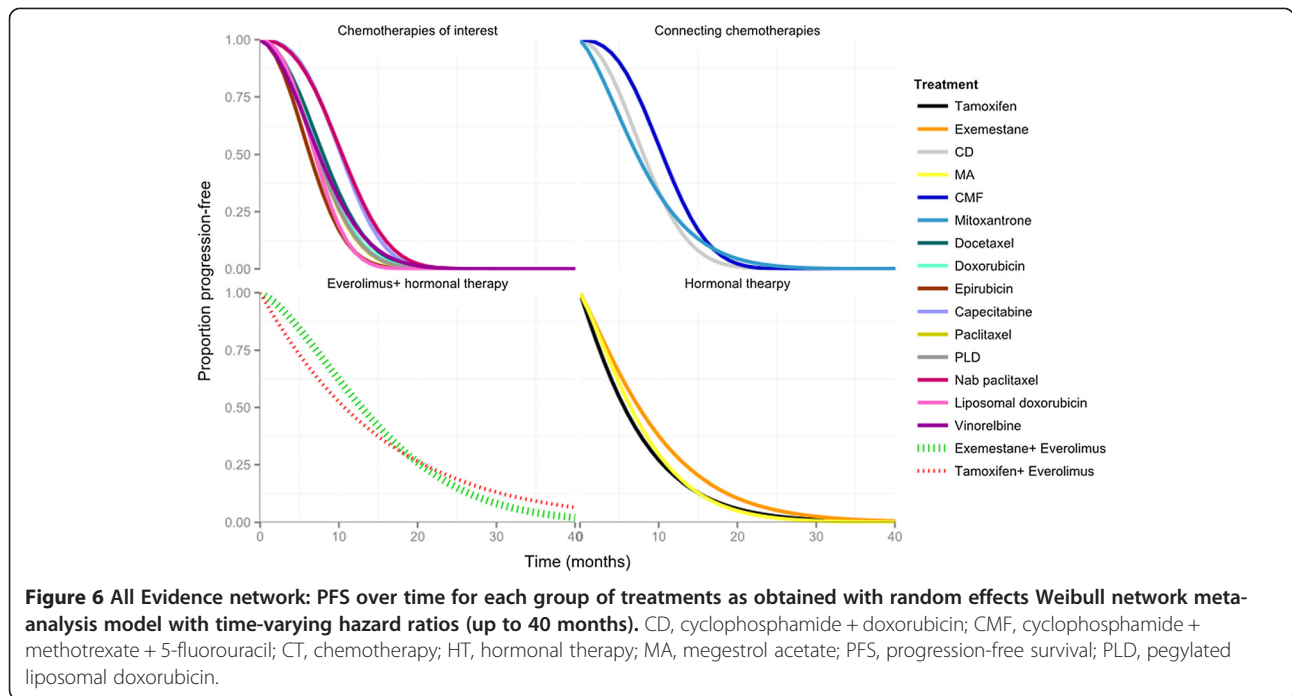
confirmed by the improved model fit of the random effects model over the fixed effects model based on the DIC (FE = 3729.0; RE = 3684.0). These results confirm the presence of heterogeneity in the network.

Given that there are three 'closed loops' within the network, an unrelated means (UM) analysis [4] was planned to assess whether results for each pairwise comparison were consistent with the estimates from the NMA which included both the direct and indirect evidence. However, due to the limited amount of data in the network, it was expected that there would be insufficient power to detect a significant difference between the NMA and UM models.

### **Results of the NMA**

Results are presented in terms of PFS curves over time for each intervention (Figure 6) and the PFS hazard ratios (over time) for each intervention relative to TAM [see Additional file 11: Figure S8]. Based on the PFS survival functions, the mean PFS was estimated at 20 and 40 months, respectively (Table 2). The mean PFS reflects the area under the PFS survival curve to the left of each time point up until the corresponding follow-up time point of interest. This represents a summary measure of PFS that does not require the curves to be fully extrapolated (that is, up until all patients have progressed). Additional detail regarding the interpretation of survival outcomes in the context of a NMA based on fractional polynomial models is discussed by Cope *et al.* [69]. All figures summarizing the results are presented up until 40 months, at which point a majority of the patients are expected to have progressed. Given the large number of curves, the treatments are grouped depending on the comparison: chemotherapies of interest, the connecting chemotherapies (that is, RCTs that connect hormonal therapies to chemotherapies), everolimus in combination with hormonal therapy, and hormonal therapy.

Results suggest that the hazard ratios for vinorelbine, CD, docetaxel, doxorubicin, PLD, paclitaxel, nab-paclitaxel, CMF, epirubicin, capecitabine and liposomal doxorubicin increased over time versus TAM, whereas the hazard ratios for everolimus + TAM, everolimus + EXE, EXE, MA and MZ were relatively constant over time versus TAM. Hence, a proportional hazard assumption is not valid, and, as such, results based on a constant hazard ratio model should be interpreted with caution. Based on the NMA using Weibull time-varying hazard ratio model, everolimus in combination with either EXE or TAM is expected to be at least as good as the alternative chemotherapies of interest. However, given the differences identified in terms of treatment doses, outcome definitions and potential treatment effect modifiers, there is a risk that the indirect treatment estimates are potentially biased. Adjusting for baseline risk did not provide a useful approach for



the network of RCTs given the small number of studies including TAM, and identifying potential inconsistencies in the network was challenging given the low power and the number of comparisons included in the closed loops.

**Discussion**

The aim of this study was to propose a more structured process to assess the feasibility of performing a valid NMA. The suggested procedure builds on existing recommendations for NMAs and provides more explicit

**Table 2 All evidence network: mean PFS per intervention and difference in expected PFS for everolimus versus alternatives**

Treatment	Mean PFS at 20 months*	95% CrI	Mean PFS at 40 months*	95% CrI
Tamoxifen	6.71	(5.31; 8.29)	7.05	(5.48; 9.11)
Exemestane	8.36	(4.55; 12.63)	9.04	(4.60; 16.47)
CD	7.97	(5.00; 11.91)	7.98	(5.00; 12.50)
MA	7.18	(4.51; 10.38)	7.42	(4.54; 11.85)
CMF	10.14	(5.85; 15.76)	10.23	(5.85; 20.36)
Mitoxantrone	7.71	(3.43; 13.03)	7.91	(3.43; 16.01)
Docetaxel	7.92	(4.37; 12.45)	7.93	(4.37; 13.30)
Doxorubicin	7.35	(3.96; 12.09)	7.36	(3.96; 12.80)
Epirubicin	6.16	(2.53; 12.17)	6.16	(2.53; 12.96)
Capecitabine	9.90	(5.03; 16.67)	9.96	(5.04; 22.43)
Paclitaxel	7.10	(3.78; 11.90)	7.11	(3.78; 12.50)
PLD	7.14	(3.09; 13.56)	7.15	(3.09; 15.12)
Nab paclitaxel	10.10	(5.42; 15.75)	10.15	(5.42; 18.33)
Liposomal doxorubicin	6.66	(3.12; 12.00)	6.66	(3.12; 12.38)
Vinorelbine	7.55	(2.99; 14.28)	7.60	(2.99; 17.71)
Exemestane + Everolimus	12.21	(6.21; 16.98)	14.14	(6.25; 26.72)
Tamoxifen + Everolimus	10.85	(5.16; 15.59)	13.62	(5.36; 24.79)

\*As obtained with random effects Weibull network meta-analysis model with time-varying hazard ratios and no covariates. CD, cyclophosphamide + doxorubicin; CMF, cyclophosphamide + methotrexate + fluorouracil; MA, megestrol acetate; PFS, progression-free survival; PLD, pegylated liposomal doxorubicin; 95% CrI, 95% credible interval.

guidance regarding the questions that should be answered at each step. The process is designed to be stepwise, with the initial stages focused on the clinical differences (that is, related to treatments, outcomes, study design and patients) and the later stages focused on evaluating the observed outcomes. Parts A and B involve an assessment of the clinical heterogeneity in terms of treatment, outcome, study and patient characteristics. Parts C and D involve an evaluation of the differences within and across the direct pairwise comparisons in terms of baseline risk and observed treatment effects. This means that it may be decided that an NMA is not feasible after the initial stage, without having assessed heterogeneity or inconsistency. If the decision is made to complete the full feasibility assessment, the available data should be illustrated and the underlying assumptions should be clearly stated, thereby improving the transparency and facilitating an interaction between methodologists and clinicians. While this process does not avoid the need for subjective decisions, it allows decision-makers or researchers to critically analyze each choice as well as to update an analysis using a different approach without necessarily having to repeat the entire process.

The final step of any NMA is to critically assess the findings. Recently guidance to facilitate this process has been developed, including the International Society for Pharmacoeconomics and Outcome Research 'instrument to assess the relevance and credibility of a NMA' [70], a 'reviewer's checklist' for evidence synthesis for treatment efficacy used in decision-making [71], as well guidance on 'how to use an article reporting (Grading of Recommendations and Evaluation (GRADE)) a multiple treatment comparison meta-analysis' [72]. Additionally, the GRADE process to assess meta-analyses has recently been updated by Cochrane to address the NMA more specifically. The current process for assessing the feasibility does not provide explicit guidance regarding the types of tools to be used for this process, but there seems to be a shared focus on some key principles that should be assessed, including the magnitude of the treatment effects, the uncertainty in the estimates and the risk of bias due to the quality of the RCTs as well as any differences in the distribution of treatment effect modifiers across direct treatment comparisons.

In the case study comparing everolimus to alternative chemotherapies in terms of PFS for women with ABC the feasibility of the NMA was determined to be limited. Although it was possible to achieve a connected network of RCTs for the comparisons of interest, differences were identified in terms of the treatment doses and the outcome definitions, which could be explored by excluding outlier studies. However, differences were also identified with respect to the pre-defined treatment effect modifiers as well as *post-hoc* differences in specific patient

characteristics that were not possible to explore based on the available data. Some variation in baseline risk within trials including TAM was observed, as was some heterogeneity in the treatment effects, whereas the inconsistency was challenging to assess in this network. In conclusion, given the differences identified in potential treatment effect modifiers which cannot be explored, there is a substantial risk that differences in these potential treatment effect modifiers may introduce bias, threatening the overall validity of the NMA, which reflects a limitation of the available data. Despite the limited feasibility of the case study, it was decided to perform the NMA for exploratory purposes. The point estimates from the analysis suggest that everolimus in combination with EXE or TAM is at least as efficacious as the chemotherapies of interest in terms of PFS. However, the comparison of interest is linked through several indirect treatment comparisons, which led to substantial uncertainty in the treatment estimates. We would advise caution regarding the interpretation of the results given the conclusion of the feasibility assessment.

The decision to proceed with the NMA can be criticized in light of the findings from the feasibility assessment. However, there is an immediate need for evidence from decision-makers given the context of the research question, as well as a potential long-term gap in the evidence, which suggests this NMA may provide the best available evidence. For example, findings from the current NMA may provide a more robust result based on the available evidence in comparison to a previous 'naïve chained indirect analysis' that multiplied a pooled hazard ratio for chemotherapy versus endocrine therapy (from the meta-analysis by Wilken *et al.*) by a hazard ratio for everolimus in combination with TAM versus TAM (based on the TAMRAD trial and assumed to be the same as everolimus in combination with EXE to EXE) [73]. Although there is a risk that results of the NMA will be over-interpreted, we would argue that the purpose of the feasibility assessment is to ensure that the underlying assumptions and limitations of the NMA are clearly communicated. Further, NMA results may help to quantify the between-study variability (and possibly the inconsistencies in the evidence base in some cases), thereby providing a more complete exploration of heterogeneity, which may generate further hypotheses [7]. Finally, in some cases, results of an NMA may actually help to trigger a response from clinical experts regarding the plausibility of the underlying assumptions, which may otherwise be more difficult to reveal. In general, we would advise consideration regarding the value of exploratory analyses against the risk of over-interpretation.

The case study of everolimus for women with ABC provides a unique opportunity to illustrate the challenges associated with evaluating the feasibility of a NMA given

that this new treatment reflects a step-change in clinical practice. In such cases where a new treatment introduces an additional step in the traditional treatment pathway, it may be necessary to compare the current treatment pathway (in the absence of the new treatment) with the anticipated treatment pathway (including the new treatment). When there are no trials available comparing the current treatment pathway to the anticipated treatment pathway, this often implies a comparison between the new treatment and the usual treatment used as the 'next step' in the treatment pathway. However, by definition, a new treatment that delays the next step in a treatment pathway is designed to target a less severe population. Consequently, there is an inherent risk that the patient characteristics of the RCTs available for the new treatment are not comparable to those patients in the RCTs evaluating the 'next step'. Additionally, as new treatments become more targeted based on genetic differences in receptors, it may be difficult to compare new trials evaluating a subset of patients with older trials including a full population (that may not report the receptors of interest). Despite these limitations, it may be decided to combine the direct and indirect results and to perform an NMA given the absence of evidence regarding the comparison of interest and the need for clinicians and health technology bodies such as NICE to make decisions. The tendency to perform an NMA in the context reinforces the importance of the feasibility assessment process. Moreover, this case study identifies a clear need for a new trial comparing the everolimus to chemotherapy, or a comparison of the alternative treatment pathways with and without everolimus (that is, everolimus followed by chemotherapy versus placebo followed by chemotherapy).

One of the main limitations of the case study is that overall survival was not assessed. The current study focused on PFS given the available data for everolimus at the time of the feasibility assessment. In comparison to overall survival, PFS is not susceptible to confounding by differences in subsequent treatments across the studies, although there is a risk of assessment bias with PFS. Therefore, overall survival, as well as the safety and adverse events of these agents, should be considered in addition to the results of the current NMA. Another limitation is that the current case study was based on a research question focused on the comparison of everolimus versus chemotherapy. However, the original scope of the research question as defined by NICE also included fulvestrant as a comparator of interest. A separate NMA has been performed by Bachelot *et al.* in order to address this comparison of interest among women with ER+ ABC following progression or recurrence after endocrine therapy [74]. Although ideally all of the comparisons of interest should be included in one simultaneous analysis, there is a clear justification for a

separate analysis in this case given the challenge of comparing everolimus to chemotherapy.

It should be noted that this feasibility process has some limitations. In the initial stages (parts A and B) it may not be necessary to extract the outcomes of interest from all studies, thereby improving the efficiency of the process. However, it is necessary to assess whether there is a sufficient amount of information reported regarding the outcome and its measure of uncertainty, which requires decision rules regarding the calculation of treatment differences or the estimates of uncertainty that may be particularly challenging to define *a priori* for continuous endpoints depending on the available information. Similarly, if imputation will be used to assess uncertainty measures, a threshold regarding the amount of missing information that will be permitted may be necessary. However, pre-specifying decision-rules for all possible types of endpoints, including optimal thresholds for the amount of data required for covariate analyses may be challenging. Although some research has evaluated alternative imputation methods for NMAs [75], to our knowledge alternative thresholds for missing data depending on the type of outcome requires further research.

Although the current case study was based on a complex network structure, in 'star' shaped networks, involving several trials with a common comparator (such as placebo), we would emphasize the importance of assessing whether differences in baseline risk exist and can be adjusted (part C). A plot of the difference measure versus the baseline risk is useful to help illustrate the variation in the baseline risk, as well as the relationship between the difference and baseline risk for each treatment. Even in cases where head-to-head trials are included in the network, it is possible to predict a placebo-arm on the basis of the other trials [15].

The current framework suggests a separate process for each outcome (and time point) of interest based on evidence available from RCTs identified from a systematic review regarding a comparative efficacy or safety question. However, undergoing the outlined feasibility process is expected to be very time consuming, and it may be more realistic to assess multiple outcomes in parallel, particularly when they are related to the same endpoints or underlying concepts. The case study explores the feasibility of a NMA based on a synthesis of Kaplan Meier curves; however, this process can be applied to any type of endpoints (that is, binary, continuous or rate outcomes). For binary endpoints it may be important to consider whether differences in follow-up are expected to act as a treatment effect modifier and, if so, to what extent different follow-up (or time points) can be combined. Similarly, for continuous outcomes, the range of time points at follow-up than can be considered comparable should be clearly addressed. It may also be important to consider models that

combine multiple time points (repeated measures) or outcomes identified within the systematic review, particularly in cases where the initial feasibility assessment suggests a NMA may not be feasible. For example, in the context of ABC, a multi-state model that accounts for PFS and overall survival (as well as the relationship between the outcomes) may provide more information (and possibly more precision).

Another possible extension of the current process would be to consider a broader evidence base if a network is deemed not to be feasible. Depending on how the research question was defined, it may be important to assess whether additional indirect evidence may be available by broadening the comparators of interest, although this consideration should be offset by the risk of introducing different populations in terms of the distribution of treatment effect modifiers. Similarly, it may be possible to integrate non-randomized evidence using more informative prior distributions [4] or individual patient data from RCTs [5,76-78] or non-randomized studies [79-81], which may influence the feasibility of an analysis. Furthermore, it may be possible to elicit bias distributions when there is insufficient data for a meta-regression where experts provide information regarding internal and external biases in order to adjust the study-specific treatment effect [82] as cited in [20]. However, these methods to combine multiple time points, outcomes and study designs are evolving currently and require further research. Therefore, the current process may provide a useful starting point to identify the need for a more complicated approach.

## Conclusions

In conclusion, the process outlined to assess the feasibility of a NMA provides a stepwise framework that will help to ensure that the underlying assumptions are systematically explored and that the risks (and benefits) of pooling and indirectly comparing treatment effects from RCTs for a particular research question are transparent.

## Additional files

**Additional file 1: Search Strategy.**

**Additional file 2: Table S1.** Trial names, publications and interventions assessed.

**Additional file 3: Figure S1.** Network of included RCTs for the base case progression-free survival analysis based on Kaplan Meier curves.

**Additional file 4: Table S2.** Treatment doses and schedules for the RCTs included in the network.

**Additional file 5: Figure S2.** Network of included RCTs for the base case PFS based on Kaplan Meier curves: risk of bias summary.

**Additional file 6: Figure S3.** Network of included RCTs for the base case PFS based on Kaplan Meier curves: risk of bias per trial.

**Additional file 7: Figure S4.** Network of included RCTs for the base case PFS based on Kaplan Meier curves: prior hormonal therapy.

**Additional file 8: Figure S5.** Network of included RCTs for the base case PFS based on Kaplan Meier curves: prior chemotherapy.

**Additional file 9: Figure S6.** PFS as extracted from Kaplan Meier curves for individual randomized controlled trials included by treatment.

**Additional file 10: Figure S7.** PFS as extracted from Kaplan Meier curves by individual randomized controlled trials included by study.

**Additional file 11: Figure S8.** All Evidence network: progression-free survival hazard ratio over time for each treatment group relative to tamoxifen as obtained with random effects Weibull network meta-analysis model with time varying hazard ratios and no covariates (up to 40 months).

## Abbreviations

ABC: advanced breast cancer; AI: aromatase inhibitor; CD: cyclophosphamide + doxorubicin; CMF: cyclophosphamide + methotrexate + fluorouracil; CT: chemotherapy; DIC: deviance information criterion; EOCG: European Co-operative Oncology Group; ER+: estrogen receptor positive; EVE: everolimus; EXE: exemestane; HER2-: human epidermal growth factor receptor 2 negative; HR-status: hormone receptor status; HT: hormonal therapy; MA: megestrol acetate; MZ: mitoxantrone (mitoxantrone); NICE: National Institute for Health and Clinical Excellence; NMA: network meta-analysis; NSA: non-steroidal aromatase inhibitors; PFS: progression-free survival; PLD: pegylated liposomal doxorubicin; RCT: randomized controlled trial; TAM: tamoxifen; TTF: time to failure; TTP: time to progression; UM: unrelated mean; 95% CrI: 95% credible interval.

## Competing interests

SC, and BS are employees of Mapi and received funding from Novartis for the study. JJ is a former employee of Mapi and received funding from Novartis for the study. JZ and SS are full time employees of Novartis Pharmaceutical Corp and have shares in the company. PS received no compensation for working on this manuscript. PS has no competing interests.

## Authors' contributions

All authors participated in the development of this manuscript. SC, JJ and BS participated in all stages of the study design, systematic literature view, statistical analyses and manuscript development. JZ, SS, and PS participated in the study design and coordination and helped to draft the manuscript. All authors read and approved the final manuscript.

## Author details

<sup>1</sup>Mapi, 33 Bloor Street East, Suite 1300, Toronto, Ontario M4W 3H1, Canada. <sup>2</sup>Novartis Pharmaceuticals Corporation, One Health Plaza, BLDG 337, A10.4C, East Hanover, NJ 07936, USA. <sup>3</sup>Mapi, 180 Canal Street, Suite 503, Boston, MA 02114, USA. <sup>4</sup>Tufts University School of Medicine, 145 Harrison Ave, Boston, MA 02111, USA. <sup>5</sup>Barts Cancer Institute, Queen Mary University of London, Old Anatomy Building, Charterhouse Square, London EC1M6BQ, UK.

Received: 7 January 2014 Accepted: 12 May 2014

Published: 5 June 2014

## References

1. Bafeta A, Trinquart L, Seror R, Ravaud P: Analysis of the systematic reviews process in reports of network meta-analyses: methodological systematic review. *BMJ* 2013, **347**:f3675.
2. National Institute for Health and Care Excellence: **Guide to the methods of technology appraisal**. 2013. [http://publications.nice.org.uk/pmg9] Published 4 April 2013; Accessed 29 July 2013.
3. Cooper NJ, Sutton AJ, Morris D, Ades AE, Welton NJ: Addressing between-study heterogeneity and inconsistency in mixed treatment comparisons: application to stroke prevention treatments in individuals with non-rheumatic atrial fibrillation. *Stat Med* 2009, **28**:1861-1881.
4. Dias S, Sutton AJ, Ades AE, Welton NJ: Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Med Decis Making* 2013, **33**:607-617.

5. Jansen JP, Fleurence R, Devine B, Iltzler R, Barrett A, Hawkins N, Lee K, Boersma C, Annemans L, Cappelleri J: **Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1.** *Value Health* 2011, **14**:429–437.
6. Jansen JP, Naci H: **Is network meta-analysis as valid as standard pairwise meta-analysis? It all depends on the distribution of effect modifiers.** *BMC Med* 2013, **11**:159.
7. Gagnier JJ, Morgenstern H, Altman DG, Berlin J, Chang S, McCulloch P, Sun X, Moher D: **Consensus-based recommendations for investigating clinical heterogeneity in systematic reviews.** *BMC Med Res Methodol* 2013, **13**:106.
8. Ades AE: **A chain of evidence with mixed comparisons: models for multi-parameter synthesis and consistency of evidence.** *Stat Med* 2003, **22**:2995–3016.
9. Bucher HC, Guyatt GH, Griffith LE, Walter SD: **The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials.** *J Clin Epidemiol* 1997, **50**:683–691.
10. Caldwell DM, Ades AE, Higgins JP: **Simultaneous comparison of multiple treatments: combining direct and indirect evidence.** *BMJ* 2005, **331**:897–900.
11. Hoaglin DC, Hawkins N, Jansen JP, Scott DA, Iltzler R, Cappelleri JC, Boersma C, Thompson D, Larholt KM, Diaz M, Barrett A: **Conducting indirect-treatment-comparison and network-meta-analysis studies: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 2.** *Value Health* 2011, **14**:429–437.
12. Lu G, Ades AE: **Combination of direct and indirect evidence in mixed treatment comparisons.** *Stat Med* 2004, **23**:3105–3124.
13. Lumley T: **Network meta-analysis for indirect treatment comparisons.** *Stat Med* 2002, **21**:2313–2324.
14. Salanti G, Higgins JPT, Ades AE, Ioannidis JPA: **Evaluation of networks of randomized trials.** *Stat Methods Med Res* 2008, **17**:279–301.
15. Achana FA, Cooper NJ, Dias S, Lu G, Rice SJ, Kendrick D, Sutton AJ: **Extending methods for investigating the relationship between treatment effect and baseline risk from pairwise meta-analysis to network meta-analysis.** *Stat Med* 2013, **32**:752–771.
16. Dias S, Welton NJ, Caldwell DM, Ades AE: **Checking consistency in mixed treatment comparison meta-analysis.** *Stat Med* 2010, **29**:932–944.
17. Lu G, Ades AE: **Assessing evidence inconsistency in mixed treatment comparisons.** *J Am Stat Assoc* 2006, **101**:447–4459.
18. White IF, Barrett JK, Jackson D, Higgins JP: **Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression.** *Res Synth Meth* 2012, **3**:111–125.
19. Higgins JP, Jackson D, Barrett JK, Lu G, Ades AE, White IR: **Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies.** *Res Synth Meth* 2012, **3**:98–110.
20. Dias S, Sutton AJ, Welton NJ, Ades AE: **Evidence synthesis for decision making 3: heterogeneity—subgroups, meta-regression, bias, and bias-adjustment.** *Med Decis Making* 2013, **33**:618.
21. Greenland S, Morgenstern H: **Ecological bias, confounding, and effect modification.** *Int J Epidemiol* 1989, **18**:269–274.
22. Berlin JA, Santanna J, Schmid CH, Szczech LA, Fedlman HI, Anti-Lymphocyte Antibody Induction Therapy Study Group: **Individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head.** *Stat Med* 2002, **21**:371–387.
23. **Afinitor Summary of Product Characteristics.** [http://www.medicines.org.uk/emc/medicine/22281/SPC] Accessed 7 October 2013.
24. Baselga J, Campone M, Piccart M, Burris HA III, Rugo HS, Sahnoud T, Noguchi S, Gnant M, Pritchard KI, Lebrun F, Beck JT, Ito Y, Yardley D, Deleu I, Perez A, Bachelot T, Vittori L, Xu Z, Cahana A, Taran T, Sahnoud T, Lebwahl D, Hortobagyi GN: **Everolimus in postmenopausal hormone-receptor-positive advanced breast cancer.** *N Engl J Med* 2012, **366**:520–529.
25. Beaver JA, Park BH: **The BOLERO-2 trial: the addition of everolimus to exemestane in the treatment of postmenopausal hormone receptor-positive advanced breast cancer.** *Future Oncol* 2012, **8**:651–657.
26. Piccart M, Noguchi S, Pritchard KI, Burris HA, Rugo HS, Gnant M, Hortobagyi GN, Melichar B, Petrakova K, Arena F, Xu C, Cahana A, Taran T, Sahnoud T, Lebwahl D, Campone M, Baselga J: **Everolimus for postmenopausal women with advanced breast cancer: updated results of the BOLERO-2 Phase III Trial.** *J Clin Oncol* 2012, **30**:suppl; abstr 559.
27. Yardley DA, Noguchi S, Pritchard KI, Burris HA III, Baselga J, Gnant M, Hortobagyi GN, Campone M, Pistilli B, Piccart M, Melichar B, Petrakova K, Arena FP, Erdkamp F, Harb WA, Feng W, Cahana A, Taran T, Lebwahl D, Rugo HS: **Everolimus plus exemestane in postmenopausal patients with HR+ Breast Cancer: BOLERO-2 final progression-free survival analysis.** *Adv Ther* 2013, **30**:870–874.
28. Bachelot T, Bourcier C, Cropet C, Ray-Coquard I, Ferrero JM, Freyer G, Abadie-Lacourtoisie S, Eymard JC, Debled M, Spaeth D, Legouffe E, Allouache D, El Kouri C, Pujade-Lauraine E: **Randomized phase II trial of everolimus in combination with tamoxifen in patients with hormone receptor-positive, human epidermal growth factor receptor 2-negative metastatic breast cancer with prior exposure to aromatase inhibitors: a GINECO study.** *J Clin Oncol* 2012, **30**:2718–2724.
29. National Institute for Health and Clinical Excellence: **Early and Locally Advanced Breast Cancer Diagnosis and Treatment.** In *NICE Clinical Guidelines, No. 80.* Cardiff (UK): National Collaborating Centre for Cancer (UK); 2009. ISBN-13: 978-0-9558265-2-8. Available at: <http://www.nice.org.uk/nicemedia/pdf/cg80niceguideline.pdf>. Accessed 7 October 2013.
30. National Institute for Health and Clinical Excellence: **Everolimus in combination with exemestane for treating advanced HER2 negative hormone-receptor-positive breast cancer after endocrine therapy.** Appraisal consultation document; March 2013. Available at: <http://guidance.nice.org.uk/TAG/327/FAD>. Accessed 7 October 2013.
31. National Institute for Health and Clinical Excellence: **Everolimus in combination with an aromatase inhibitor for the treatment of advanced or metastatic HER2 negative, oestrogen receptor positive breast cancer after prior endocrine therapy, Final scope.** Single Technology Appraisal; October 2012. Available at: <http://www.nice.org.uk/nicemedia/live/13866/61357/61357.pdf>. Accessed 7 October 2013.
32. **Advanced Breast Cancer: Endocrine Therapy.** [http://pathways.nice.org.uk/pathways/advanced-breast-cancer#path=view%3A/pathways/advanced-breast-cancer/advanced-breast-cancer-endocrine-therapy.xml&content=close].
33. Jansen JP, Cope S: **Meta-regression models to address heterogeneity and inconsistency in network meta-analysis of survival outcomes.** *BMC Med Res Methodol* 2012, **12**:152.
34. Saad ED, Katz A: **Progression-free survival and time to progression as primary end points in advanced breast cancer: often used, sometimes loosely defined.** *Ann Oncol* 2009, **20**:460–464.
35. Ouwens MJ, Philips Z, Jansen JP: **Network meta-analysis of parametric survival curves.** *Res Synth Meth* 2011, **1**:258–271.
36. Jansen JP: **Network meta-analysis of survival data with fractional polynomials.** *BMC Med Res Methodol* 2011, **11**:61.
37. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A: **Bayesian measures of model complexity and fit.** *J R Stat Soc (Series B)* 2002, **64**:583–639.
38. Dempster AP: **The direct use of likelihood for significance testing.** *Stat Comput* 1997, **7**:247–252.
39. Cope S, Ouwens MJ, Jansen JP, Schmid P: **Progression-free survival with fulvestrant 500 mg and alternative endocrine therapies as second-line treatment for advanced breast cancer: a network meta-analysis with parametric survival models.** *Value Health* 2013, **16**:403–417.
40. **A randomized trial in postmenopausal patients with advanced breast cancer comparing endocrine and cytotoxic therapy given sequentially or in combination.** The Australian and New Zealand Breast Cancer Trials Group, Clinical Oncological Society of Australia. *J Clin Oncol* 1986, **4**:186–193.
41. Beuselink B, Wildiers H, Wynendael W, Dirix L, Kains JP, Paridaens R: **Weekly paclitaxel versus weekly docetaxel in elderly or frail patients with metastatic breast carcinoma: a randomized phase-II study of the Belgian Society of Medical Oncology.** *Crit Rev Oncol Hematol* 2010, **75**:70–77.
42. Bontenbal M, Andersson M, Wildiers J, Cocconi G, Jassem J, Paridaens R, Rotmensz N, Sylvester R, Mouridsen HT, Klijn JG, van Oosterom AT: **Doxorubicin vs. epirubicin, report of a second-line randomized phase II/III study in advanced breast cancer.** EORTC Breast Cancer Cooperative Group. *Br J Cancer* 1998, **77**:2257–2263.
43. Dixon AR, Jackson L, Chan S, Haybittle J, Blamey RW: **A randomised trial of second-line hormone vs single agent chemotherapy in tamoxifen resistant advanced breast cancer.** *Br J Cancer* 1992, **66**:402–404.
44. Gill PG, Gebiski V, Snyder R, Burns I, Levi J, Byrne M, Coates A: **Randomized comparison of the effects of tamoxifen, megestrol acetate, or tamoxifen plus megestrol acetate on treatment response and survival in patients with metastatic breast cancer.** *Ann Oncol* 1993, **4**:741–744.

45. Gradishar WJ, Krasnojon D, Cheporov S, Makhson AN, Manikhas GM, Clawson A, Bhar P: **Significantly longer progression-free survival with nab-paclitaxel compared with docetaxel as first-line therapy for metastatic breast cancer.** *J Clin Oncol* 2009, **27**:3611–3619.
46. Gradishar WJ, Krasnojon D, Cheporov S, Makhson AN, Manikhas GM, Clawson A, Bhar P, McGuire JR, Iglesias J: **Phase II trial of nab-paclitaxel compared with docetaxel as first-line chemotherapy in patients with metastatic breast cancer: final analysis of overall survival.** *Clin Breast Cancer* 2012, **12**:313–321.
47. Henderson IC, Allegra JC, Woodcock T, Wolff S, Bryan S, Cartwright K, Dukart G, Henry D: **Randomized clinical trial comparing mitoxantrone with doxorubicin in previously treated patients with metastatic breast cancer.** *J Clin Oncol* 1989, **7**:560–571.
48. Ingle JN, Ahmann DL, Green SJ, Edmonson JH, Creagan ET, Hahn RG, Rubin J: **Randomized clinical trial of megestrol acetate versus tamoxifen in parmenopausal or castrated women with advanced breast cancer.** *Am J Clin Oncol* 1982, **5**:155–160.
49. Jones SE, Erban J, Overmoyer B, Budd GT, Hutchins L, Lower E, Laufman L, Sundaram S, Urba WJ, Pritchard KI, Mennel R, Richards D, Olsen S, Meyers ML, Ravdin PM: **Randomized phase III study of docetaxel compared with paclitaxel in metastatic breast cancer.** *J Clin Oncol* 2005, **23**:5542–5551.
50. Katsumata N, Watanabe T, Minami H, Aogi K, Tabei T, Sano M, Masuda N, Andoh J, Ikeda T, Shibata T, Takashima S: **Phase III trial of doxorubicin plus cyclophosphamide (AC), docetaxel, and alternating AC and docetaxel as front-line chemotherapy for metastatic breast cancer: Japan clinical oncology group trial (JCOG9802).** *Ann Oncol* 2009, **20**:1210–1215.
51. Kaufmann M, Bajetta E, Dirix LY, Fein LE, Jones SE, Zilembo N, Dugardyn JL, Nasurdi C, Mennel RG, Cervek J, Fowst C, Polli A, di Salle E, Arkhipov A, Piscitelli G, Miller LL, Massimini G: **Exemestane is superior to megestrol acetate after tamoxifen failure in postmenopausal women with advanced breast cancer: results of a phase III randomized double-blind trial.** *J Clin Oncol* 2000, **18**:1399–1411.
52. Kaufmann M, Bajetta E, Dirix LY, Fein LE, Jones SE, Zilembo N, Dugardyn JL, Nasurdi C, Mennel RG, Cervek J, Fowst C, Polli A, di Salle E, Arkhipov A, Piscitelli G, Miller LL, Massimini G: **Exemestane improves survival in metastatic breast cancer: results of a phase III randomized study.** *Clin Breast Cancer* 2000, **1**:515–518.
53. Kaufmann M, Bajetta E, Dirix LY, Fein LE, Jones SE, Cervek J, Fowst C, Polli A, Di Salle E, Massimini G, Piscitelli G: **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.** *Eur J Cancer* 2000, **36**:586–587.
54. Meier CR, Illiger HJ, Steder M, Janssen J, Deertz H, Braun M, Oeney HT, Deuss B, Kuchler T, Rotermund S: **Weekly vinorelbine versus docetaxel for metastatic breast cancer after failing anthracycline treatment.** *Onkologie* 2008, **31**:447–453.
55. Muss HB, Paschold EH, Black WR, Cooper MR, Capizzi RL, Christian R, Cruz JM, Jackson DV, Stuart JJ, Richards F, White DR, Zekan PJ, Spurr CL, Pope E, Cose D, Morgan T, Wells HB: **Megestrol acetate v tamoxifen in advanced breast cancer: a phase III trial of the Piedmont Oncology Association (POA).** *Semin Oncol* 1985, **12**:55–61.
56. Muss HB, Wells HB, Paschold EH, Black WR, Cooper MR, Capizzi RL, Christian R, Cruz JM, Jackson DV, Powell BL: **Megestrol acetate versus tamoxifen in advanced breast cancer: 5-year analysis—a phase III trial of the Piedmont Oncology Association.** *J Clin Oncol* 1988, **6**:1098–1106.
57. O'Brien ME, Wigler N, Inbar M, Rosso R, Grischke E, Santoro A, Catane R, Kieback DG, Tomczak P, Ackland SP, Orlandi F, Mellars L, Alland L, Tendler C: **Reduced cardiotoxicity and comparable efficacy in a phase III trial of pegylated liposomal doxorubicin HCl (CAELYX/Doxil) versus conventional doxorubicin for first-line treatment of metastatic breast cancer.** *Ann Oncol* 2004, **15**:440–449.
58. O'Shaughnessy JA, Blum J, Moiseyenko V, Jones SE, Miles D, Bell D, Rosso R, Mauriac L, Osterwalder B, Burger HU, Laws S: **Randomized, open-label, phase II trial of oral capecitabine (Xeloda) vs. a reference arm of intravenous CMF (cyclophosphamide, methotrexate and 5-fluorouracil) as first-line therapy for advanced/metastatic breast cancer.** *Ann Oncol* 2001, **12**:1247–1254.
59. Paridaens R, Biganzoli L, Bruening P, Klijn JG, Gamucci T, Houston S, Coleman R, Schachter J, Van Vreckem A, Sylvester R, Awada A, Wildiers J, Piccart M: **Paclitaxel versus doxorubicin as first-line single-agent chemotherapy for metastatic breast cancer: a European Organization for Research and Treatment of Cancer Randomized Study with cross-over.** *J Clin Oncol* 2000, **18**:724–733.
60. Paridaens R, Dirix L, Beex L, Nooij M, Cufer T, Lohrisch C, Biganzoli L, Hoorebeeck I, Duchateau L, Lobelle JP, Piccart M: **Promising results with exemestane in the first-line treatment of metastatic breast cancer: a randomized phase II EORTC trial with a tamoxifen control.** *Clin Breast Cancer* 2000, **1**:S19–S21.
61. Paridaens R, Dirix L, Beex L, Nooij M, Cufer T, Lohrisch C, Biganzoli L, Hoorebeeck I, Duchateau L, Lobelle JP, Piccart M: **Mature results of a randomized phase II multicenter study of exemestane versus tamoxifen as first-line hormone therapy for postmenopausal women with metastatic breast cancer.** *Ann Oncol* 2003, **14**:1391–1398.
62. Paridaens RJ, Dirix LY, Beex LV, Nooij M, Cameron DA, Cufer T, Piccart MJ, Bogaerts J, Therasse P: **Phase III study comparing exemestane with tamoxifen as first-line hormonal treatment of metastatic breast cancer in postmenopausal women: the European Organisation for Research and Treatment of Cancer Breast Cancer Cooperative Group.** *J Clin Oncol* 2008, **26**:4883–4890.
63. Sledge GW, Neuberg D, Bernardo P, Ingle JN, Martino S, Rowinsky EK, Wood WC: **Phase III trial of doxorubicin, paclitaxel, and the combination of doxorubicin and paclitaxel as front-line chemotherapy for metastatic breast cancer: an intergroup trial (E1193).** *J Clin Oncol* 2003, **21**:588–592.
64. Villalon AH, Ngangel CA, Eufemio G, Laudico AV: **Cyclophosphamide methotrexate 5-fluorouracil (CMF), vs. CMF megestrol acetate (CMF-M), vs. Megestrol acetate (M) for advanced breast cancer in Filipino women.** *Philipp J Intern Med* 1993, **31**:115–126.
65. Chan S, Friedrich K, Noel D, Pinter T, Van Belle S, Vorobiof D, Duarte R, Gil M, Bodrogi I, Murray E, Yelle L, von Minckwitz G, Korec S, Simmonds P, Buzzi F, Mancha RG, Richardson G, Walpole E, Ronzoni M, Murawsky M, Alaki M, Riva A, Crown J: **Prospective randomized trial of docetaxel versus doxorubicin in patients with metastatic breast cancer.** *J Clin Oncol* 1999, **17**:2341–2354.
66. Yardley DA, Burris HA III, Spigel DR, Clark BL, Vazquez E, Shipley D, Barton J, Thompson D, Montes I, Greco FA, Hainsworth JD: **A phase II randomized crossover study of liposomal doxorubicin versus weekly docetaxel in the first-line treatment of women with metastatic breast cancer.** *Clin Breast Cancer* 2009, **9**:247–252.
67. Cooper NJ, Peters J, Lai MC, Juni P, Wandel S, Palmer S, Paulden M, Conti S, Welton NJ, Abrams KR, Bujkiewicz S, Spiegelhalter D, Sutton AJ: **How valuable are multiple treatment comparison methods in evidence-based health-care evaluation?** *Value Health* 2011, **14**:371–380.
68. Cowan JD, Neidhart J, McClure S, Coltman CA, Gumbart C, Martino S, Hutchins LF, Stephens RL, Vaughan CB, Osborne CK: **Randomized trial of doxorubicin, bisantrene and mitoxantrone in advanced breast cancer: a Southwest Oncology Group Study.** *J Natl Cancer Inst* 1991, **83**:1077–1084.
69. Cope S, Jansen JP: **Quantitative summaries of treatment effect estimates obtained with network meta-analysis of survival curves to inform decision-making.** *BMC Med Res Methodol* 2013, **13**:147.
70. Jansen JP, Trikalinos T, Cappelleri JC, Daw J, Andes S, Eldessouki R, Salanti G: **Indirect treatment comparison/ network meta-analysis study questionnaire to assess relevance and credibility to inform health care decision-making: an ISPOR-AMCP-NPC good practice task force report.** *Value Health* 2014, **17**:157–73.
71. Ades AE, Caldwell DM, Reken S, Welton NJ, Sutton AJ, Dias S: **Evidence synthesis for decision making 7: a reviewer's checklist.** *Med Decis Making* 2013, **33**:679–691.
72. Mills EJ, Ioannidis JP, Thorlund K, Schünemann HJ, Puhan MA, Guyatt GH: **How to use an article reporting a multiple treatment comparison meta-analysis.** *JAMA* 2012, **308**:1246–1253.
73. Fleeman N, Bagust A, Beale S, Blundell M, Dwan K, Pilkington G, Proudlove C, Dunder Y, Vecchio F, Thorp N: **Everolimus in combination with an aromatase inhibitor for the treatment of breast cancer after prior endocrine therapy: a single technology appraisal.** *LRIg, The University of Liverpool* 2013, [http://www.nice.org.uk/nicemedia/live/13866/63219/63219.pdf] Accessed 7 Oct 2013.
74. Bachelot T, McCool R, Duffy S, Glanville J, Varley D, Fleetwood K, Zhang J, Jerusalem G: **Comparative efficacy of everolimus plus exemestane versus fulvestrant for hormone-receptor-positive advanced breast cancer following progression/recurrence after endocrine therapy: a network meta-analysis.** *Breast Cancer Res Treat* 2014, **143**:125–133.
75. Spinelli LM, Higgins JP, Cipriani A, Leucht S, Salanti G: **Evaluating the impact of imputations for missing participant outcome data in a network meta-analysis.** *Clin Trials* 2013, **10**:378–388.



76. Cope S, Capkun-Niggli G, Gale R, Lassen C, Owen R, Ouwens MJ, Bergman G, Jansen JP: **Efficacy of once-daily indacaterol relative to alternative bronchodilators in COPD: a patient-level mixed treatment comparison.** *Value Health* 2012, **15**:524–533.
77. Donegan S, Williamson P, D'Alessandro U, Tudur Smith C: **Assessing the consistency assumption by exploring treatment by covariate interactions in mixed treatment comparison meta-analysis: individual patient-level covariates versus aggregate trial-level covariates.** *Stat Med* 2012, **31**:3840–3857.
78. Saramago P, Sutton AJ, Cooper NJ, Manca A: **Mixed treatment comparisons using aggregate and individual participant level data.** *Stat Med* 2012, **31**:3516–3536.
79. McCarron CE, Pullenayegum EM, Thabane L, Goeree R, Tarride JE: **Bayesian hierarchical models combining different study types and adjusting for covariate imbalances: a simulation study to assess model performance.** *PLoS One* 2011, **6**:e25635.
80. McCarron CE, Pullenayegum EM, Thabane L, Goeree R, Tarride JE: **The impact of using informative priors in a Bayesian cost-effectiveness analysis: an application of endovascular versus open surgical repair for abdominal aortic aneurysms in high-risk patients.** *Med Decis Making* 2013, **33**:437–450.
81. Schmitz S, Adams R, Walsh C: **Incorporating data from various trial designs into a mixed treatment comparison model.** *Stat Med* 2013, **32**:2935–2949.
82. Turner RM, Spiegelhalter DJ, Smith GC, Thompson SG: **Bias modelling in evidence synthesis.** *J R Stat Soc Ser A* 2009, **172**:21–47.

doi:10.1186/1741-7015-12-93

**Cite this article as:** Cope *et al.*: A process for assessing the feasibility of a network meta-analysis: a case study of everolimus in combination with hormonal therapy versus chemotherapy for advanced breast cancer. *BMC Medicine* 2014 **12**:93.

**Submit your next manuscript to BioMed Central  
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

