T-DM1 for treating HER2-positive, unresectable locally advanced or metastatic breast cancer after treatment with trastuzumab and a taxane

Produced by School of Health and Related Research (ScHARR), The University of Sheffield

Authors Hazel Squires, ScHARR, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA
Emma Simpson, ScHARR, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA
Rebecca Harvey, ScHARR, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA
John Stevens, ScHARR, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA
Helen Buckley Woods, ScHARR, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA
Matt Stevenson, ScHARR, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA

Correspondence to Hazel Squires, ScHARR, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA

Date completed 25/02/2014

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 12/65/01STA

Declared competing interests of the authors

None.

Acknowledgements

The authors wish to thank Professor Rob Coleman, Professor of Oncology, Department of Oncology Cancer Clinical Trials Centre, Academic Unit of Clinical Oncology, Weston Park Hospital, Sheffield, and Lynda Wyld, Consultant and Senior Lecturer, Department of Oncology, The Medical School, Sheffield, for providing clinical advice and commenting on draft materials during the project. We would also like to thank Gill Rooney, Programme Administrator, ScHARR, for her help in preparing and formatting the report.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Squires H, Simpson EL, Harvey R, Stevens J, Woods HB, Stevenson M. T-DM1 for treating HER2-positive, unresectable locally advanced or metastatic breast cancer after treatment with trastuzumab and a taxane: A Single Technology Appraisal. ScHARR, The University of Sheffield, 2014.

Contributions of authors

Hazel Squires acted as project lead and health economic modeller on this assessment and critiqued the manufacturer's economic evaluation. Emma Simpson critiqued the clinical effectiveness methods and evidence. Rebecca Harvey and John Stevens critiqued the statistical analyses included in the manufacturer's submission. Helen Buckley Woods commented on the searches included in the manufacturer's submission. Matt Stevenson provided health economic expertise. All authors contributed to the writing of the report.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

LIST OF ABBREVIATIONS

AE	Adverse event
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
BC	Breast cancer
BCV	Base case value
BIOSIS	Biosciences Information Service
BOLERO-3	Trial name
CDF	Cancer Drugs Fund
CEREBEL	Trial name
CI	Confidence Interval
CODA	Convergence Diagnostic and Output Analysis
DSU	Decision Support Unit
ECHO	Echocardiogram
ECOG	Eastern Cooperative Oncology Group
ECOG PS	Eastern Cooperative Oncology Group performance status scale
EMA	European Agency for the Evaluation of Medicinal Products
EMILIA	Trial name
EPAR	European Public Assessment Report
EQ-5D	EuroQol 5D
ER	Oestrogen ('Estrogen') Receptor (status)
ER+ / ER-	Oestrogen receptor positive or negative status
ERG	Evidence review group
FACT-B	Functional Assessment of Cancer Therapy-for patients with Breast Cancer
FACT-B TOI-PFB	Functional Assessment of Cancer Therapy-for patients with Breast Cancer Trial Outcomes Index-Physical/Functional/Breast
HER2	Human epidermal growth factor receptor 2
HER2-positive	Overexpression of HER2
HR	Hazard Ratio
HRQoL	Health related quality of life
ICER	Incremental cost-effectiveness ratio
LABC	Locally advanced breast cancer
LVEF	Left Ventricular Ejection Fraction
LY	Life Year
MBC	Metastatic breast cancer
MeSH	Medical subject headings

MS	Manufacturer submission to NICE
MTC	Mixed treatment comparison
MUGA scan	Multiple Gated Acquisition scan
NA	Not applicable
NE	Not estimable
NICE	National Institute for Health and Care Excellence
OS	Overall survival
PFS	Progression free survival
PR	Progesterone Receptor (status)
PRO	Patient Reported Outcome
QALY	Quality adjusted life year
RCT	Randomised controlled trial
RECIST	Response evaluation criteria in solid tumours
SAE	Serious adverse event
SE	Standard Error
T-DM1	T-DM1
TH3RESA	Trial name
TPC	Treatment of physician's choice

1. SUMMARY

1.1 Critique of the decision problem in the manufacturer's submission

The decision problem addressed by the manufacturer's submission (MS) was in line with the final scope issued by the National Institute for Health and Care Excellence (NICE).

The target population was people with overexpression of the human epidermal growth factor receptor 2 (HER2-positive), unresectable locally advanced or metastatic breast cancer whose disease has progressed after treatment with trastuzumab and a taxane. The intervention was trastuzumab emtansine (T-DM1) within its licensed indication. The following comparators were all considered within the manufacturer's submission: lapatinib in combination with capecitabine; capecitabine; vinorelbine; trastuzumab in combination with capecitabine; and trastuzumab in combination with vinorelbine. Progression-free survival (PFS) and overall survival (OS) were considered separately within a mixed treatment comparison (MTC). Adverse effects of treatment were considered only within the narrative synthesis and little is described within the manufacturer's submission in relation to health-related quality of life. The health economic outcome employed was the incremental cost per quality-adjusted life year (QALY) gained, as set out within the NICE reference case.

The description of the decision problem within the manufacturer's submission did not highlight any equity issues and there is currently no Patient Access Scheme application.

1.2 Summary of clinical effectiveness evidence submitted by the manufacturer

The clinical effectiveness data relevant to the decision problem were taken from two Randomised Controlled Trials (RCTs) of T-DM1, with adverse event (AE) data taken from a pooled analysis of additional trials of T-DM1 as a single agent (i.e. not in combination with other agents). Data from these two RCTs reported a significant advantage in PFS for T-DM1 over lapatinib in combination with capecitabine, and over the treatment of physician's choice. Data reported a significant advantage in OS and time to symptom worsening for T-DM1 over lapatinib in combination with capecitabine. The most common grade 3 or greater AEs for T-DM1 were thrombocytopenia and hepatotoxicity.

The only head-to-head RCT data were for T-DM1 compared with lapatinib in combination with capecitabine. The manufacturer also submitted a MTC which provided hazard ratios for T-DM1 versus: lapatinib in combination with capecitabine; trastuzumab in combination with capecitabine; and capecitabine monotherapy. Most of the data were from third-line or later therapy, whereas the MS suggested T-DM1 as second-line treatment.

1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

The ERG believes that all trials relevant to the decision problem with available data were included within the MS. The clinical effectiveness data relevant to the decision problem were taken from two large RCTs, both of which were open-label, but otherwise at low risk of bias.

Within the MTC, T-DM1 was the best treatment in terms of both OS and PFS. Allowing for heterogeneity between studies increased the uncertainty about the true treatment effect on OS and PFS. From the ERG's random effects model, T-DM1 was associated with a reduction in the hazard of death of 32% and in the hazard of progression or death of 35% compared to lapatinib in combination with capecitabine (the next best option).

1.4 Summary of cost effectiveness submitted evidence by the manufacturer

The manufacturer identified no existing economic evaluations of T-DM1. A *de novo* cohort state transition model was developed which adheres to the NICE Reference Case. The model has three health states: progression-free survival; post-progression; and death, and follows weekly cycles. The model was based upon the EMILIA trial comparing T-DM1 with lapatinib in combination with capecitabine. The trial data was extrapolated (with a range of approaches being tested within sensitivity analyses) and hazard ratios were applied for all other comparators based upon the MTC. A utility was assigned to each health state according to a published mixed model analysis. Costs applied to the health states included: the treatment options; their administration; treatment of a selection of AEs; supportive care; and treatment within the post-progression state.

Following the clarification process, the manufacturer's reported a deterministic ICER for T-DM1 compared with lapatinib in combination with capecitabine of £167,253, the latter of which was estimated to have an ICER of £39,449 compared with capecitabine monotherapy. All other comparators were dominated (less effectiveness with the same or higher cost, or more costly with the same or lower effectiveness) than these treatment options.

1.5 Summary of the ERG's critique of cost effectiveness evidence submitted

The *de novo* model developed is appropriate for the decision problem defined in the final scope and was generally well described within the report. The model structure was considered by the ERG to be clinically appropriate. The ERG identified two key errors in implementation and four key assumptions which were methodologically weak which were revised for the ERG's base case. However, this produced a very similar revised base case Incremental Cost-Effectiveness Ratio (ICER) to the manufacturer's of £166,429, since not all changes acted upon the ICER in the same direction.

The uncertainty around the model inputs for the probabilistic sensitivity analysis (PSA) was inappropriately characterised within the MS. The one way sensitivity analysis provided by the manufacturer did not establish the robustness of the model results or determine the key drivers of the results because T-DM1 was compared with capecitabine only.

1.6 ERG commentary on the robustness of evidence submitted by the manufacturer

1.6.1 Strengths

The key strengths of the clinical evidence were that effectiveness data were available from large clinical trials and these were RCTs, mostly at low risk of bias. In addition, no relevant RCTs were excluded from the review. Additional adverse event data were available from studies of single-agent T-DM1.

The health economic model submitted by the manufacturer was clinically appropriate and generally well described and justified.

1.6.2 Weaknesses and areas of uncertainty

The effectiveness data were from open-label trials. The lack of blinding introduces bias, especially for patient reported outcomes. Few patients contributing data were on second-line therapy, whilst the manufacturer suggests that they would anticipate T-DM1 being provided second-line. In addition, few patients had Eastern Cooperative Oncology Group performance status scale 2, whilst in practice this would constitute around one third of patients.

There is a lack of direct head-to-head comparisons with all but one comparator (lapatinib in combination with capecitabine) from the NICE final scope. The MTC submitted by the manufacturer provided hazard ratios for capecitabine monotherapy and trastuzumab in combination with capecitabine; however no evidence was identified for vinorelbine monotherapy or trastuzumab in combination with vinorelbine.

There is uncertainty around the long term PFS and OS, and this impacts substantially upon the health economic model results. In addition, the PSA submitted by the manufacturer inadequately characterised the uncertainty around the model inputs.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

The ERG produced a revised deterministic base case which was very similar to the manufacturer's base case following the clarification process. The cost per QALY gained for T-DM1 compared with lapatinib in combination with capecitabine was estimated to be £166,429, with the latter having an

ICER of £50,620 compared with capecitabine monotherapy. All other comparators were dominated by these treatment options.

The deterministic univariate sensitivity analysis undertaken by the ERG suggested that the key drivers of the model results are: the relative OS associated with the interventions; the distribution employed for extrapolation of PFS and OS; whether the treatment effect is assumed to continue beyond the trial data; the utility values associated with PFS and post-progression; and whether wastage is included within the drug costs. However, the ICER for T-DM1 versus lapatinib in combination with capecitabine did not decrease below £147,000 within any of the univariate sensitivity analyses.

2. BACKGROUND

2.1 Critique of manufacturer's description of underlying health problem.

The manufacturer's submission (MS) clearly describes HER2-positive breast cancer, defined as breast cancers leading to the activation of multiple signalling pathways within the cells resulting in an increase in their proliferation and a reduction in cell death. The MS states that unresectable locally advanced breast cancer and metastatic breast cancer remain largely incurable, with the majority of people dying due to their disease. They also highlight that without targeted therapy, HER2-positive metastatic breast cancer is associated with aggressive disease, higher rates of recurrence, shorter disease-free survival and shorter overall survival as compared with tumours that do not overexpress HER2.

The number of people eligible for treatment is based upon analyses commissioned by the manufacturer, some details of which are provided in supplementary files within the MS. The estimate includes only those people with metastatic HER2-positive breast cancer in England rather than those people with metastatic or locally advanced HER2-positive breast cancer in England and Wales, which would lead to an underestimate of eligible patients. However, our clinical experts suggest that the proportion of patients with metastatic disease being HER2-positive may be overestimated because the use of adjuvant trastuzumab has decreased incidence of metastatic disease in this patient group. Thus, whilst the number of people eligible for treatment provided within the MS appears reasonable, there is some uncertainty around the estimate.

2.2 Critique of manufacturer's overview of current service provision

The manufacturer outlines the treatment pathways specified within the latest National Institute for Health and Care Excellence (NICE) clinical guidelines,[1] but suggests that due to the introduction of the Cancer Drugs Fund (CDF) this does not reflect current service provision. An alternative treatment pathway approved by the CDF is reported. This is shown in Table 1. Clinical advice received by the Evidence Review Group (ERG) suggests that this description of current service provision is appropriate and relevant to the decision problem under consideration. However, the submission describes only the first three lines of therapy for metastatic breast cancer; yet the included trials and the ERG's clinical advisors suggest that patients may receive subsequent lines of therapy, with later lines consisting of capecitabine, vinorelbine or trastuzumab.

Table 1: Clinical pathway described within the MS

Treatment line	NICE approved clinical pathway	CDF approved clinical pathway
First-line	Trastuzumab plus paclitaxel	Pertuzumab plus trastuzumab plus docetaxel; or trastuzumab plus taxane
Second-line	Capecitabine or vinorelbine (plus trastuzumab in central nervous system only progression)	Lapatinib plus capecitabine
Third-line	Vinorelbine or capecitabine or trastuzumab	Vinorelbine or capecitabine or trastuzumab

3. CRITIQUE OF MANUFACTURER'S DEFINITION OF DECISION PROBLEM

3.1 Population

The patient population addressed by the manufacturer's statement of the decision problem matches that described in the final NICE scope. The patient population is people with HER2-positive, unresectable locally advanced or metastatic breast cancer whose disease has progressed after treatment with trastuzumab and a taxane (paclitaxel or docetaxel). In reviewing the decision problem, the manufacturer has not restricted population by race, gender or geographical location, which is consistent with the final NICE scope. In line with the licensed indication of trastuzumab emtansine (T-DM1), only adult patients are eligible for treatment. Clinical evidence was available on this population, which reflects the characteristics of the patient population in England and Wales that is eligible for treatment.

3.2 Intervention

The intervention addressed by the manufacturer's statement of the decision problem matches that described in the final NICE scope. The intervention is T-DM1 within its licensed indication. T-DM1 as a single agent is indicated for the treatment of adult patients with HER2-positive, unresectable locally advanced or metastatic breast cancer who previously received trastuzumab and a taxane, separately or in combination.[2] Patients should have either:

- Received prior therapy for locally advanced or metastatic disease; or
- Developed disease recurrence during or within six months of completing adjuvant therapy.

3.3 Comparators

The final NICE scope lists the following as comparators, which are all considered within the MS:

- lapatinib in combination with capecitabine;
- capecitabine;
- vinorelbine;
- trastuzumab in combination with capecitabine;
- trastuzumab in combination with vinorelbine.

The MTC does not include vinorelbine or trastuzumab in combination with vinorelbine because insufficient evidence is available within RCTs for these comparators. As a result, estimates of effectiveness for vinorelbine monotherapy and trastuzumab in combination with vinorelbine within the health economic model are assumed to be the same as those for capecitabine monotherapy and trastuzumab in combination with capecitabine respectively. This assumption is based upon NICE

Clinical Guidelines 81 where the guideline development group agreed that the effectiveness of the two treatments is essentially equivalent.[1]

3.4 Outcomes

The final NICE scope lists the following outcome measures, all of which were considered within the MS:

- progression free survival (PFS);
- overall survival (OS);
- adverse effects of treatment;
- health-related quality of life (HRQoL).

Only PFS and OS were considered within the mixed treatment comparison (MTC). Adverse effects of treatment were considered only within the narrative synthesis and little is described within the MS around HRQoL.

The health economic outcome employed is the incremental cost per quality-adjusted life year (QALY) gained, as set out within the NICE reference case.

3.5 Other relevant factors

The description of the decision problem within the MS does not highlight any equity issues. However, within the clinical effectiveness section, the manufacturer highlights that the patient population is younger on average than the general breast cancer population and has particularly aggressive disease. They suggest that this leads to increased broader societal impacts of the disease including effects on family life, as well as personal and societal financial implications.

There is currently no Patient Access Scheme application.

The MS highlights that pertuzumab as a first-line therapy within this patient population is approved by the CDF and is currently being reviewed by NICE. They suggest that there is currently insufficient safety data around the use of T-DM1 following pertuzumab, but that prospective studies are planned to evaluate the safety of T-DM1 after pertuzumab. The use of pertuzumab may have implications for the effectiveness and cost-effectiveness of T-DM1; however this is beyond the scope of this assessment.

4. CLINICAL EFFECTIVENESS

4.1 Critique of the methods of review(s)

Section 6 of the MS consists of three components for clinical evidence:

A) A systematic review of the decision problem (Sections 6.1 and 6.2 and Appendix 10.2 of the MS, discussed in Sections 4.1.1 – 4.1.5 of ERG report). This review identified T-DM1 trials of relevance to the final scope from NICE.

B) A systematic review to populate the MTC (Section 6.7 of the MS and Appendix 10.4 of the MS, discussed in Section 4.1.6 of ERG report). This review attempted to identify data to allow T-DM1 to be compared against all the comparators listed in the final NICE scope.

C) Adverse event data, for which no further systematic review was conducted, but trials additional to those from the systematic review of the decision problem were included (Section 6.9 of the MS, discussed in Section 4.1.7 of ERG report).

4.1.1 Searches

A search was conducted to identify "all randomised evidence relevant to the decision problem". Keywords and subject headings for breast cancer or metastatic breast cancer were combined with terms for the intervention (T-DM1). This search was then limited by the use of terms to indicate the study type: randomised controlled trial. Medline, Embase, Embase Alert, BIOSIS and the Cochrane Library were the primary data sources used. In addition a number of conference websites were searched as well as an internal data source "PubCentre". Although various conference proceedings were searched, searches of trials registers, for example via the Current Controlled Trials website and a broader source such as the Science Citation Index, were not searched. These could have been searched as a safeguard that no studies had been missed. It is not possible to comment on the validity of the PubCentre database as this is an internal resource of the manufacturer. No supplementary techniques such as citation or reference searching were reported.

In this search strategy, it would have been preferable to expand the free text terms used to express the problem concepts, as over-reliance on subject headings to identify evidence is not the best methodological approach; especially for new technologies or interventions. In Embase, the terms pro 132365, pro132365, t dm 1, t dm1, tmab mcc dm1 and trastuzumab dm1 are not subject headings (as suggested in the manufacturer's clarifications) but a list to show the user that if they wish to search for one of these terms they should use T-DM1/. Using these as free text terms would potentially have increased the sensitivity of the search. This is important when there will be no further search

iterations. Similarly the terms for the study type could have been expanded by using a published sensitive study filter. In The Cochrane library it is possible to use Medical subject headings (MeSH) and also to combine search statements. These features of the database could have been utilised in order to apply the search as had been done in the other data sources (e.g. Medline).

Despite these shortcomings the searches were believed satisfactory to retrieve all the relevant evidence that the ERG and clinical advisors are aware of and, given the recent nature of the intervention, it is unlikely that any relevant studies have been missed for the clinical effectiveness review.

4.1.2 Inclusion criteria: Systematic review of decision problem.

Section 6.2 of the MS describes study selection for the systematic review of the decision problem. The inclusion and exclusion criteria were not complete in this section, but were presumed to be supplementary to the defined decision problem in Section 5 of the MS. Those inclusion and exclusion criteria explicit from Section 6.2 of the MS were consistent with the scope of the decision problem.

Population

The population evaluated was metastatic breast cancer (MBC), or patients with unresectable locally advanced breast cancer (LABC), limited to studies of humans. This was consistent with the scope of the decision problem. Section 6.2 of the MS did not specify that cancer had to be HER2-positive, although that is specified in Section 5 of the MS which describes the decision problem. The delivery of T-DM1 within the licensed indication would also mean that only HER2-positive cancer would apply. No inclusion or exclusion criteria are listed relating to race, gender or geographical location. Section 6.2 of the MS does not indicate any restriction in age of population. However, Section 5 of the MS specifies that the decision problem is restricted to T-DM1 within the licensed indication, which would mean only adult patients were eligible for treatment.

Intervention

Studies of T-DM1 as a single agent were included. T-DM1 in combination with other agents was excluded. This was consistent with the scope of the decision problem. Although not specified in Section 6.2, Section 5 of the MS specifies that the decision problem is restricted to T-DM1 within its licensed indication. T-DM1 as a single agent is indicated for the treatment of adult patients with HER2-positive, unresectable locally advanced or metastatic breast cancer who previously received trastuzumab and a taxane, separately or in combination.[2] Patients should have either received prior therapy for locally advanced or metastatic disease, or developed disease recurrence during or within six months of completing adjuvant therapy.

Comparators

Included and excluded comparators were not explicit from Section 6.2.1 of the MS. However, Section 5 of the MS specifies that the comparators are as per the final NICE scope, namely: lapatinib in combination with capecitabine; capecitabine; vinorelbine; trastuzumab in combination with capecitabine; and trastuzumab in combination with vinorelbine. Section 6.2.6 of the MS excludes a study (TDM4450g)[3] that has a comparator outside the NICE scope, although the comparator is not the only reason given for exclusion of this study, as only a small proportion of patients received T-DM1 within its licensed indication.

Outcomes

Included and excluded outcomes were not explicit from Section 6.2.1 of the MS. Section 5 of the MS specifies that the outcomes are as per the final NICE scope, namely: PFS; OS; AEs; HRQoL.

Study design

RCTs were included, and other study types were excluded. This is appropriate given that there are RCTs addressing the decision problem.

It was unclear how many reviewers conducted study selection. However the study selection process was explicit from Appendix 10.2.3 of the MS, with reasons for study exclusion listed.

4.1.3 Data extraction for the systematic review of the decision problem.

The method of data extraction was not explicit from either section 6.2 or Appendix 10.2 of the MS. It is unclear how many reviewers were involved, or whether pre-specified questions were addressed. However, adequate details of both included trials (EMILIA and TH3RESA[4,5]) were included in Section 6 of the MS. Details of trial characteristics and outcomes were accurate, as checked against published data. Not all HRQoL outcomes were reported. However, a reference was given to published data from EMILIA.[6]

4.1.4 Quality assessment for the systematic review of decision problem.

Section 6.4 of the MS provided a quality assessment of the included trials within the systematic review. It is unclear how many reviewers were involved. Tables 11–13 of the MS provided information on both included RCTs, the EMILIA[4] and TH3RESA[5] trials. The quality criteria addressed were taken from the Centre for Reviews and Dissemination[7], and were appropriately chosen given that the included trials were both RCTs.

Quality assessment of the studies was accurate, with the following possible exceptions. For the TH3RESA[5] trial, in answer to the question "Was the concealment of treatment allocation adequate?", the MS state "No". However, this appears to refer to lack of blinding of the study. The employment of central allocation, via Interactive Voice/Web Response System, implies that allocation to either intervention or comparator arm would be adequately concealed, that is, not known in advance of assignment. For the TH3RESA trial, in answer to the question "Were there any unexpected imbalances in drop-outs between groups?", it is stated that a greater number of patients in the Treatment of Physician's Choice (TPC) arm (13.1%) than in the T-DM1 arm (4.7%) decided to withdraw from the study, the MS answers "No". This may be the case, however given the differences in numbers, it should be made explicit whether attrition bias was avoided, by stating whether the remaining participants in the two treatment groups were still balanced in terms of prognostic factors. As an ITT analysis is provided, this should avoid bias.

4.1.5 Evidence synthesis for the systematic review of decision problem.

No meta-analysis was conducted of the trials included from the systematic review of the decision problem. This was appropriate given that a MTC was conducted (discussed in Sections 4.3 and 4.4 of the ERG report).

4.1.6 Systematic review to populate the MTC.

4.1.6.1 Searches: Systematic review to population the MTC

A search was conducted to identify RCTs and non-RCTs (to enhance completeness) for a MTC. Medline, Medline in Process, Embase and The Cochrane Library were the data sources used. A number of conference websites were also searched for additional evidence. The search has a well-developed vocabulary to reflect the concepts of the decision problem and the relevant study types. The search concepts were mapped against the population under investigation as defined by the EMILIA study in order to create greater specificity in the search results. The search was limited from 1998 (the date of approval of trastuzumab) to December 2012. For completeness, there was scope to re-run the search with the same date range as the clinical effectiveness search (which was run in October 2013). Aside from hand searching of conference websites no other supplementary searching was reported.

4.1.6.2 Inclusion criteria: Systematic review to populate the MTC

Section 6.7.3 and Appendix 10.4 of the MS describes inclusion and exclusion criteria of the systematic review to populate the MTC.

Population

The population included was people with unresectable HER2-positive LABC or MBC that progressed after previous treatment with trastuzumab and a taxane in the adjuvant or metastatic setting. Progression had to occur during or after the most recent treatment for LABC or MBC or within six months after treatment for early-stage disease. The population was limited by age to those aged 18 years or over, with no restriction on gender or race. This was appropriate for the MTC, and matched the scope of the decision problem. The target population for study selection was described as "rigorously defined EMILIA-matched population criteria", meaning that the populations would be within licensed indications for T-DM1, even though other treatments would not necessarily have those same patient population restrictions on their licensed indications in practice. This matched the scope of the decision problem.

Intervention and comparators

All pharmacological interventions for treatment of HER2-positive unresectable LABC or MBC were included. This was appropriate in attempting to build a network for the MTC.

Outcomes

Included and excluded outcomes were not explicit from Section 6.7.3 of the MS. However, from Section 6.7.5, it was apparent that only OS and PFS were considered.

Study design

RCTs or non-RCTs were included, not restricted by phase of study, or whether the study was blinded or not.

Section 6.7.3 of the MS states that study selection was made by two reviewers, with a third reviewer used to resolve any disagreements. This is good practice for a systematic review.

4.1.6.3 Data extraction: Systematic review to populate MTC

The method of data extraction was given in Appendix 10.4 of the MS. Two reviewers extracted data independently, with involvement of a third reviewer where necessary to resolve disputes. This is good practice for a systematic review. Adequate details are given of included trials in Section 6.7 of the MS.

4.1.6.4 Quality assessment: Systematic review to populate MTC

Section 6.7.3 of the MS describes the critical appraisal process for the systematic review to populate the MTC. Criteria for the appraisal of RCTs was as for the systematic review of the decision problem, that is, based on the quality criteria adapted from the Centre for Reviews and Dissemination.[7] A quality assessment strategy for non-randomised studies had been planned. However, the non-randomised studies identified from the search were excluded from the MTC due to not linking into the network.

4.1.6.5 Evidence synthesis: Systematic review to populate MTC

See MTC critique section (Section 4.4 of ERG report).

4.1.7 Adverse event data

No further review was conducted, but adverse event data was presented from trials additional to those included in the systematic review of the decision problem (Section 6.9 of the MS). Evidence was presented from the pooled analysis from the European Public Assessment Report (EPAR).[2] An update of this analysis was provided in the manufacturer's response to clarification questions. Adverse event data were provided from an abstract of an additional study, and adverse event data from both of the RCTs included in the systematic review of the decision problem were reported.

4.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

The ERG believe that that there were no unidentified RCTs with available data meeting the inclusion criteria in the final NICE scope. Trials of T-DM1 in HER2-positive LABC or MBC are listed in Appendix 1. There were no inappropriate exclusions. One RCT was excluded from the systematic review of the decision problem at full text sift, (trial TDM4450g), as T-DM1 was not prescribed within its licensed indication for most patients (no prior MBC treatment), and the comparator was outside the scope (trastuzumab and docetaxel). A search of clinicaltrials.gov[8] identified 33 trials of T-DM1 in breast cancer, of which 16 were not mentioned in the MS. However, none of these trials were comparative trials relevant to the decision problem.

4.2.1 Clinical effectiveness trials included in the review

Effectiveness data were taken from two phase III RCTs, the EMILIA[4,6] and TH3RESA trials.[5] Table 2 presents characteristics of trials reported in the MS that contained effectiveness data.

Table 2: Characteristics of trials with effectiveness data reported in MS

Trial identifier(s)	Number of patients	Population	Intervention	Comparator	Outcomes
EMILIA; NCT00829166; TDM4370g; BO21977	991	LABC or MBC, HER2-positive Prior treatment for breast cancer in the adjuvant, unresectable, locally advanced, or metastatic setting must include both a taxane, alone or in combination with another agent, and trastuzumab, alone or in combination with another agent	TDM 1 3.6mg/kg intravenously every 21 days	lapatinib 1250 mg/day orally once per day of each 21-day cycle plus capecitabine 1000 mg/m^2 orally twice daily on Days 1-14 of each 21-day treatment cycle	OS, PFS, Objective Response, Duration of Objective Response, Clinical benefit (the percentage of patients with a complete response, partial response, or stable disease at 6 months after randomisation), Time to Treatment Failure, Time to symptom progression (defined as the time from randomisation to the first symptom progression as measured by FACT-B) Adverse events
TH3RESA; NCT01419197; TDM4997g ; BO25734	602	LABC or MBC, HER2-positive Prior treatment with an trastuzumab, a taxane, and lapatinib, disease progression after at least two regimens of HER2-directed therapy in the metastatic or unresectable locally advanced/recurrent setting	TDM 1 3.6mg/kg intravenously every 21 days	Treatment of Physician's Choice (chemotherapy, hormonal therapy, biologic drug and/or HER2-directed therapy)	OS, PFS, Objective response rate, Duration of objective response, Land mark survival rate (6 months/1 year), Time to pain symptom progression as measured by the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire BM22, Global Health Status/Quality of Life as measured by the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire C30 , Global Health Status as measured by Euro-Qol 5D Adverse events

Main clinical effectiveness trials

OS and PFS data were taken from two open-label RCTs, the EMILIA[4] and TH3RESA[5] trials. Both were international, multi-centre studies, with centres in Europe (including the UK), the United States and Asia. Both trials investigated T-DM1 within its licensed indication in populations of HER2-positive LABC or MBC. In the EMILIA trial, the comparator was lapatinib in combination with capecitabine. In the TH3RESA[5] trial, the comparator was treatment of physician's choice (TPC), of which: 68.5% of patients received chemotherapy plus trastuzumab; 10.3% of patients received lapatinib plus trastuzumab; 1.6% of patients received hormonal therapy plus trastuzumab; 2.7% of patients received chemotherapy plus lapatinib; and 16.8% of patients received single-agent chemotherapy.[5]

At the time of the MS submission, both the EMILIA and TH3RESA trials were ongoing. Both trials had completed the primary endpoint and effectiveness data were available.

For EMILIA, the primary analysis took place with a clinical data cut-off of 14 January 2012. Following the primary PFS analysis, a formal request was received from regulatory authorities for an additional analysis of OS prior to the planned protocol-specified final analysis. This second interim analysis of OS was conducted with data cut-off date of 31 July 2012. All data in the MS and the publication of Verma *et al*[4] were prior to treatment switching. A final analysis of OS, following patient switching, is planned when 632 events are reached.

For TH3RESA[5], the primary analysis took place with a clinical data cut-off of 11 February 2012. All patients in the TPC arm were given the option of switching to the T-DM1 arm at progression. At the time of analysis 44 of the 198 patients in the TPC arm had switched over to receive T-DM1. A final analysis of OS is planned when 492 events have been observed.

Both trials were randomised; EMILIA[4] was randomised 1:1 to either T-DM1 or lapatinib in combination with capecitabine, and TH3RESA[5] was randomised 2:1 to either T-DM1 or TPC. In both trials, randomisation was stratified by world region and prognostic factors. For EMILIA[4] the factors were: world region (United States, Western Europe, other); and within each of the four categories defined by the following two prognostic factors, the number of prior chemotherapeutic regimens for unresectable, locally advanced or metastatic disease (0−1 vs. > 1), and any visceral versus no visceral disease. FOR TH3RESA[5] the factors were: world region (United States, Western Europe, other); number of prior regimens (excluding single-agent hormones) for the treatment of metastatic or locally advanced/recurrent unresectable disease (2−3 or > 3); and any visceral disease versus no visceral disease.

Eligibility criteria are listed in Appendix 2 of the ERG report. Ineligibility of screened patients in both trials was mostly due to HER2 status not matching inclusion criteria or presence of brain metastases that were untreated, symptomatic, or required therapy to control symptoms, and screening was conducted prior to randomisation, so was unlikely to introduce bias between treatment groups. Both EMILIA[4] and TH3RESA[5] trials included patients with HER2-positive unresectable LABC or MBC. Both trials required adequate organ function and Left Ventricular Ejection Fraction (LVEF) greater than or equal to 50% by either an echocardiogram (ECHO) or MUltiple Gated Acquisition (MUGA) scan. Both trials required prior trastuzumab and taxane treatment; TH3RESA additionally required prior lapatinib treatment. TH3RESA[5] required at least two prior lines of treatment for MBC or LABC, and so patients were, on average, on later lines of therapy than in the EMILIA[4] trial. EMILIA[4] restricted the population to ECOG 0-1, whereas TH3RESA[5] also allowed patients with ECOG 2 to be included.

Baseline characteristics from the trials are shown in Table 3. (Table adapted from the MS and trial publications.[6])

At baseline, in the EMILIA[4] trial considering lines of therapy defining prior systemic therapy as any systemic endocrine or chemotherapy, 12% of patients were first-line, 36% of patients were second-line and 52% of patients were third or later lines. (If considering only chemotherapeutic regimens, 39% of patients had had more than one prior chemotherapy regimen for LABC or MBC). In the TH3RESA[5] trial, 35% of patients were third-line, 36% of patients were fourth-line and 29% of patients were fifth or later lines. The studies were international, so not all treatment choices would have been consistent with the UK clinical pathway. Most patients were female. There were five male patients in the EMILIA [4] trial; one in the T-DM1 arm and four in the comparator arm.

Table 3: Baseline characteristics of EMILIA [4] and TH3RESA[5] trials

Baseline characteristic	EMILIA		TH3RESA	
	T-DM1 (n = 495)	Lapatinib plus capecitabine (n = 496)	T-DM1 (n = 404)	TPC (n = 198)
Race (%)				
White	72	75	80.4	81.3
Asian	19	17	14.1	12.1
Black/African American	6	4	-	-
Other	1	2	5.4	6.6
Not available	1	1		
World region (%)				
United States	27	27	24.5	24.2
Western Europe	32	32	42.3	42.9
Asia	17	15	-	-
Other	25	25	33.2	32.8
Median age, y (range)	53 (25–84)	53 (24–83)		
Age, %				
<65 years			85.4	82.8
65–74 years			11.4	14.1
≥75 years			3.2	3.0
ECOG PS 0, (%)	60	63	44.8	41.4
ECOG PS 1, (%)	39	35	49.8	51.0
ECOG PS 2, (%)	0	0	5.5	7.6
Not available	1	2		
Measurable disease by independent review, n (%)	397 (80)	389 (78)		
Metastatic involvement, (%)				
Visceral	67	68	74.8	75.8
Non-visceral	33	32	25.2	24.2
Metastatic sites, (%)				
<3	57	62		
≥3	41	35		
Unknown	2	3		
Brain metastasis at baseline, %	9	10	9.9	13.6
Disease extent at study entry, %				

Baseline characteristic	**EMILIA**		**TH3RESA**	
	T-DM1 (n = 495)	**Lapatinib plus capecitabine (n = 496)**	**T-DM1 (n = 404)**	**TPC (n = 198)**
Metastatic Unresectable locally advanced/recurrent BC			96.8 3.2	94.4 5.6
Number of prior regimens for LABC/MBC, median (range)			4 (1–14)	4 (1–19)
Number of prior regimens for LABC/MBC 0-1 >1	 61 39	 61 39	NA	NA
Number of prior regimens for LABC/MBC ≤3, % 4–5, % >5, %	 NA	 NA	 32.6 37.1 30.3	 39.4 32.8 27.8
Prior trastuzumab treatment For MBC or both early and MBC For early BC only	 84 16	 84 16		
ER/PR status (%) ER+ and/or PR+ ER- and PR- Unknown	 57 41 2	 53 45 2	51.5	52.0

Quality assessment showed the trials were at low risk of bias, apart from the lack of blinding, as shown within Table 4.

Lack of blinding was stated in the MS as being due to the number of placebo treatments that would have been needed for the control arm, and "obvious drug effects in at least some patients". An imbalance in frequency of physician visits between arms can introduce bias. However, for both trials, tumour assessments were conducted approximately every 6 weeks. Lack of blinding of patients and physicians is likely to have introduced bias in the trials, especially for the HRQoL outcomes. This is less likely to affect OS results. PFS may be prone to bias in unblinded studies. EMILIA [4] , but not TH3RESA[5], had blinded outcome assessment for PFS to address this source of bias.

Table 4: Quality assessment of clinical effectiveness trials

Question	**EMILIA**	**TH3RESA**
Was randomisation carried out appropriately?	Yes	Yes
Was the concealment of treatment allocation adequate?	Yes	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Yes	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation?	No (except blinding of outcome assessors for PFS)	No
Were there any unexpected imbalances in drop-outs between groups?	No	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No
Did the analysis include an intention-to-treat analysis?	Yes	Yes

4.2.2 Clinical effectiveness trials Overall survival data

Table 5 shows the OS data reported from the EMILIA[4] and TH3RESA[5] trials. As the patients from the TH3RESA trials were at a later stage in treatment on average, it is unsurprising that median OS is lower in TH3RESA than for EMILIA.

Table 5: Overall survival data from clinical effectiveness trials

Results	**EMILIA**		**TH3RESA**	
	T-DM1	**Lapatinib plus capecitabine**	**T-DM1**	**TPC**
OS median (months)	30.9	25.1	14.9	NE (not estimable)
Stratified HR	0.682 (95% CI 0.55-0.85) p=0.0006		0.552 (95% CI 0.369-0.826) P=0.0034	

CI: Confidence Interval

The EMILIA[4] results in Table 5 were based on a second interim OS analysis (data cut-off July 31st 2012), with all data prior to patients switching treatment. The first interim OS analysis of EMILIA [4] had a hazard ratio (HR) of 0.621 (95% CI: 0.475 - 0.813). There were landmark analyses at 1 and 2 years showing 85.2% patients and 78.4% alive at 1 year and 64.7% patients and 51.8% patients alive at 2 years with T-DM1 and lapatinib in combination with capecitabine, respectively.

The TH3RESA[5] results in Table 5 were based on interim analysis of 105 events (21% of targeted events). This was statistically non-significant i.e. not meeting the stopping boundary, and further OS analyses for the trial were planned, but not conducted at the time of writing.

Tables 6 shows HRs for subgroups for which randomisation was stratified.

Table 6: Subgroup OS data from the EMILIA[4] trial

Subgroup	Subgroup	No. of patients	HR (95% CI)
No. of prior chemotherapeutic regimens for LABC or MBC	0-1	609	0.80 (0.61-1.07)
	>1	382	0.58 (0.41-0.81)
World region	United States	270	0.62 (0.41-0.96)
	Western Europe	317	0.95 (0.65-1.39)
	Asia	158	0.48 (0.27-0.85)
	Other	246	0.68 (0.45-1.04)
Disease involvement	Visceral	669	0.59 (0.46-0.76)
	Non-visceral	322	1.05 (0.69-1.61)

In the EMILIA[4] trial, randomisation was stratified within each of the four categories defined by the following two prognostic factors: the number of prior chemotherapeutic regimens for unresectable, locally advanced or metastatic disease (0−1 versus > 1); and visceral versus non-visceral disease. However, this was based on investigator assessed definitions of visceral or non-visceral. Randomisation was also stratified by world region, although it should be noted that stratification was by three categories (United States, Western Europe, other); whereas subgroup data was presented by four categories (see Table 6).

The MS provides an additional file looking at subgroups of EMILIA based on two definitions of visceral versus non-visceral disease [Data on File RXUKDONF00337, December 2013 MS]. Where visceral disease was defined as lung and liver involvement, the HR for OS for the non-visceral subgroup was 0.73 (95% CI: 0.48 - 1.12). Where visceral disease was defined as lung, liver, pleural effusion and ascites, the HR for OS for the non-visceral subgroup was 0.59 (95% CI: 0.37 - 0.94).

Other subgroup OS analyses, shown in Appendix 3 of the ERG report, are not stratified. Most subgroups favoured T-DM1 over the comparator, although not all reached statistical significance. For patients aged 75 years and over, OS results favoured the comparator over T-DM1, HR 3.45 (95% CI: 0.94 - 12.65). This was based on data from 25 patients, and median survival for the comparator was not estimable. For patients with brain metastases (not requiring therapy to control symptoms) there was a treatment group difference for OS favouring T-DM1 (n=45) over lapatinib and capecitabine (n=50), HR 0.382 (95% CI: 0.184-0.795). No subgroup analyses were presented for the interim OS analysis of TH3RESA.[5]

4.2.3 Clinical effectiveness trials Progression free survival data

Both EMILIA [4] and TH3RESA[5] defined PFS as survival free from death or progression, based on the Response Evaluation Criteria in Solid Tumours (RECIST) method of tumour response (MS p69). EMILIA [4] had a primary outcome of PFS by independent review, but also measured investigator assessed PFS. TH3RESA[5] had a primary outcome of investigator assessed PFS.

Table 7 shows EMILIA [4] PFS by independent review (HR= 0.650). Investigator assessed PFS in the EMILIA trials was also statistically significant, HR=0.658 (95% CI: 0.56-0.78) p<0.001. Table 7 shows TH3RESA[5] ITT analysis of primary endpoint in which 44 patients from TPC had switched to T-DM1.

Table 7: Progression free survival from clinical effectiveness trials

Results	**EMILIA**		**TH3RESA**	
	T-DM1 N=495	**Lapatinib plus capecitabine N=496**	**T-DM1 N=404**	**TPC N=198**
PFS median months	9.6	6.4	6.2	3.3
Stratified HR	0.650 (95% CI 0.55-0.77) P<0.0001		0.528 (95% CI 0.422-0.661) P<0.0001	

Table 8 shows HRs for subgroups in EMILIA [4] and TH3RESA[5] for which randomisation was stratified. In the EMILIA [4] study randomisation was stratified within each of the four categories defined by the following two prognostic factors, the number of prior chemotherapeutic regimens for unresectable, locally advanced or metastatic disease (0−1 versus > 1), and visceral vs non-visceral disease. However, this was based on investigator assessed definitions of visceral or non-visceral. Randomisation was also stratified by world region, although it should be noted that stratification was by the three categories United States, Western Europe, other; whereas subgroup data was presented by four categories (see Table 8).

In the TH3RESA[5] study, randomisation was stratified by world region (United States, Western Europe, or Other) and presence of visceral disease by investigator assessed definitions (any visceral disease versus no visceral disease). It was also stratified by number of prior regimens (excluding single-agent hormones) for the treatment of metastatic or locally advanced/recurrent unresectable

disease (2−3 or > 3), although it should be noted there is a further subdivision in results presented for more than six prior regimens.

Table 8: PFS for subgroups (unstratified HRs)

Subgroup	**Subgroup**	**No. of patients EMILIA**	**EMILIA HR (95% CI)**	**No. of patients TH3RESA**	**TH3RESA HR (95%CI)**
No. of prior chemotherapeutic regimens for LABC or MBC	0-1	609	0.68 (0.55-0.58)	-	-
	>1	382	0.63 (0.49-0.82)	-	-
No. of prior regimens (excluding single-agent hormones) for LABC or MBC	≤3	-	-	209	0.48 (0.32-0.70)
	4-6	-	-	214	0.58 (0.40-0.83)
	>6	-	-	177	0.48 (0.32-0.73)
World region	United States	270	0.70 (0.51-0.98)	147	0.71 (0.44-1.14)
	Western Europe	317	0.56 (0.41-0.74)	256	0.44 (0.32-0.61)
	Asia	158	0.74 (0.50-1.08)	-	-
	Other	246	0.73 (0.51-1.03)	199	0.53 (0.36-0.78)
Disease involvement	Visceral	669	0.55 (0.45-0.67)	452	0.56 (0.44-0.72)
	Non-visceral	322	0.96 (0.71-1.30)	150	0.41 (0.26-0.64)

Treatment effects of T-DM1 versus lapatinib in combination with capecitabine were less certain for non-visceral disease, and there were too few patients over 75 to draw conclusions on this age group.
The MS provides an additional file looking at subgroups of EMILIA based on two definitions of visceral versus non-visceral disease [Data on File RXUKDONF00337, December 2013 MS]. Where visceral disease was defined as lung and liver involvement, the HR for IRC-assessed PFS for the non-visceral subgroup was 0.76 (95% CI: 0.56 - 1.02). Where visceral disease was defined as lung, liver,

pleural effusion and ascites, the HR for the independent review committee assessed PFS for the non-visceral subgroup was 0.69 (95% CI: 0.51 - 0.95).

For the TH3RESA[5] trial, with investigator defined visceral versus non-visceral disease, the median PFS for patients with visceral disease was 3.4 months in the TPC arm compared to 6.2 months in the T-DM1 arm (HR=0.56 (95% CI: 0.44 - 0.72)), and for patients with non-visceral disease was 3.1 months in the TPC arm compared to 6.7 months in the T-DM1 arm (HR=0.41 (95% CI: 0.26 - 0.64)).

Other subgroup data, shown in Appendix 3 of the ERG report, are not stratified. Most subgroups favoured T-DM1 over the comparator, although not all reached statistical significance. In the EMILIA trial, for patients aged 75 and over PFS results favoured the comparator over T-DM1, HR 3.51 (95% CI: 1.22 - 10.13). This was based on data from 25 patients, a sample size too small to draw conclusions. For patients with brain metastases in the EMILIA trial (n=95), there was no statistically significant treatment group difference for PFS. For TH3RESA[5], results were not presented separately for each treatment of physician's choice in the comparator arm within the MS or available publications, except for considering trastuzumab containing regimens. For the TH3RESA[5] trial, if the comparator was limited to treatment regimens including trastuzumab (n=149), the median survival was 3.2 months, HR=0.558 (95% CI: 0.437 - 0.711) compared with T-DM1 ($p<0.0001$), which was similar to the results with the whole TPC group analysed as the comparator.

4.2.4 Clinical effectiveness trials HRQoL data

Final analyses of HRQoL outcomes for TH3RESA were not available at the time of writing due to the trial being ongoing (although utility scores based on interim EQ-5D measurements from TH3RESA were provided confidentially in a supplementary file with the MS [Roche DoF RXUKDONF00339 Dec 2013]).

HRQoL data were available from the EMILIA trial, which used a patient reported outcome (PRO) to assess time to symptom progression in the female study participants, which was the primary PRO endpoint. Time to symptom progression was defined as the time from randomisation to the first symptom progression as measured by the Functional Assessment of Cancer Therapy-for patients with Breast Cancer (FACT-B) questionnaire with the Trial Outcomes Index-Physical/Functional/Breast (TOI-PFB) subscale. The FACT-B TOI-PFB subscale contains 23 items from the FACT-B questionnaire: physical well-being; functional well-being; and additional concerns for breast cancer patients (breast cancer subscale). All items in the questionnaire were rated by the patient on a 5-point scale ranging from 0 (not at all) to 4 (very much), with a higher score indicating better perceived quality of life. A change of 5 points or more is considered clinically meaningful.[6,8]

Analyses were conducted on data provided by female patients with baseline and at least one post-baseline score. There were 450 T-DM1 patients and 445 lapatinib in combination with capecitabine patients in the primary PRO endpoint analysis. The median time to symptom progression, defined as a decrease from baseline of 5 points or more in the FACT-B TOI score, was statistically significantly longer in the T-DM1 group at 7.1 months, compared with 4.6 months in the lapatinib plus capecitabine group; stratified hazard ratio 0.796 (95% CI: 0.667 - 0.951, p=0.0121).[6] These data suggest that deterioration took longer in the T-DM1 group. Two sensitivity analyses of the primary PRO endpoint were conducted. When symptom worsening that occurred after missing assessments was backdated to the last non-missing assessment date plus one (to assess the effect of missing assessments), results were 6.0 months for T-DM1 versus 4.3 months for lapatinib in combination with capecitabine (stratified HR 0.788 (95% CI: 0.660 - 0.941), p=0.0089).[6] When the date of symptom worsening was backdated by six weeks (to investigate potential bias due to delayed symptom reporting), results were 6.6 months for T-DM1 versus 4.2 months for lapatinib in combination with capecitabine (stratified HR 0.820 (95% CI: 0.686 - 0.979), p=0.0286).[6]

There were two predefined exploratory PRO endpoints: the proportion of patients with a clinically significant improvement in symptoms between the two treatment arms as measured by the FACT-B TOI-PFB; and the proportion of patients with diarrhoea symptoms as measured by the four-item Diarrhoea Assessment Scale (DAS).[6] In the T-DM1 arm, 249/450 patients (55.3% (95% CI: 50.7% - 60.0%)) developed clinically meaningful improvement in symptoms from baseline compared with 220/445 patients (49.4% (95% CI: 44.7% - 54.2%)) in the lapatinib plus capecitabine arm. This was not statistically significantly different between treatment groups (p = 0.0842).[6] Although similar at baseline, the number of patients reporting diarrhoea symptoms increased 1.5- to 2-fold during treatment with lapatinib in combination with capecitabine but remained near baseline levels in the T-DM1 arm.[6]

4.2.5 Adverse events

Adverse event data was presented from the two RCTs included in the systematic review of the decision problem. In addition, evidence was presented from the pooled analysis from the EPAR,[2] and an update of this analysis was provided by the manufacturer's response to clarification questions. Adverse event data were also provided from an abstract of an additional study.[9] Table 9 shows trial characteristics of trials with safety data reported in the MS.

The pooled analysis available for the EPAR (and to the US Food and Drug Administration) had 882 patients, and is also available as a conference abstract.[10] The MS provided updated pooled analysis with 884 patients (cut-off date for pooled analysis 31.07.2012). The pooled analysis includes data

from the T-DM1 groups of the EMILIA[4] and TDM4450g trials, but does not include data from TH3RESA.[5]

Table 9: Trial characteristics of trials with safety data reported in MS

Trial identifier(s)	Number of patients in trial	Population	Intervention	Comparator	Adverse event data
EMILIA; NCT00829166; TDM4370g; BO21977	991	LABC or MBC, HER2-positive Prior treatment for breast cancer in the adjuvant, unresectable, locally advanced, or metastatic setting must include both a taxane, alone or in combination with another agent, and trastuzumab, alone or in combination with another agent	T-DM1 3.6mg/kg intravenously every 21 days	lapatinib 1250 mg/day orally once per day of each 21-day cycle plus capecitabine 1000 mg/m^2 orally twice daily on Days 1-14 of each 21-day treatment cycle	Adverse event data for trial reported (n=490 T-DM1, n=488 comparator), and also T-DM1 group data in pooled analysis (n=490 patients T-DM1 in pooled analysis)
TH3RESA; NCT01419197; TDM4997g ; BO25734	602	LABC or MBC, HER2-positive Prior treatment with an trastuzumab, a taxane, and lapatinib, disease progression after at least two regimens of HER2-directed therapy in the metastatic or unresectable locally advanced/recurrent setting	T-DM1 3.6mg/kg intravenously every 21 days	Treatment of Physician's Choice (chemotherapy, hormonal therapy, biologic drug and/or HER2-directed therapy)	Adverse event data for trial reported (n=403 T-DM1, n=198 comparator) (not in pooled analysis)
NCT00679341; TDM4450g ; BO21976	137 in trial	LABC or MBC, HER2-positive No prior chemotherapy for their MBC	T-DM1 3.6mg/kg intravenously every 21 days	loading dose of trastuzumab 8 mg/kg IV + docetaxel 75 or 100 mg/m^2 IV on Day 1 of Cycle 1 followed by trastuzumab 6 mg/kg IV + docetaxel 75 or 100 mg/m^2 IV on Day 1 of all subsequent 21-day cycles	Adverse event data in pooled analysis n=106 (Data from 69 patients randomised to T-DM1, and 37 patients who crossed over from the control arm)
NCT00943670; TDM4688g	51	LABC or MBC, HER2-positive, History of prior trastuzumab therapy	T-DM1 3.6mg/kg intravenously every 21 days From Cycle 4, participants with early progressive disease additional pertuzumab by IV infusion at a loading dose of 840 mg on Day 1, followed by 420 mg IV infusion every 3 weeks	NA	Adverse event data in pooled analysis n=51 from single agent phase of study

Trial identifier(s)	Number of patients in trial	Population	Intervention	Comparator	Adverse event data
NCT00932373; TDM3569g	55	LABC or MBC, HER2-positive, progression during or within 60 days after treatment with any prior trastuzumab-containing chemotherapy regimen Previous treatment with chemotherapy for MBC	T-DM1 various doses, including licensed dose of T-DM1 3.6mg/kg intravenously every 21 days	NA	Adverse event data in pooled analysis, from the n=15 patients on the licensed dose of T-DM1
NCT00509769; TDM4258g	112	MBC, HER2-positive Prior HER2 targeted therapy	T-DM1 3.6mg/kg every 3 weeks	NA	Adverse event data in pooled analysis n=112
NCT00679211; TDM4374g	110	MBC, HER2-positive at least 2 lines of therapy	T-DM1 3.6mg/kg every 3 weeks	NA	Adverse event data in pooled analysis n=110
NCT00781612; TDM4529g ; BO25430	145 planned	LABC or MBC, HER2-positive prior TDM	single-agent T-DM1; or combination T-DM1 administered in combination with paclitaxel or with pertuzumab ± paclitaxel	NA	Pooled analysis: followed-up patients who were in the above-listed studies and in this extension study
JO22997	73	HER2-positive MBC. prior treatment with trastuzumab and at least 1 chemotherapy	T-DM1 3.6mg/kg every 3 weeks	NA	Adverse event data available from abstract (not in pooled analysis)

Adverse events (AEs) from EMILIA [4] and TH3RESA[5] are shown in Table 10, which is taken directly from the MS.

Table 10: AEs from EMILIA and TH3RESA, table adapted from MS

System organ/ class/adverse events	EMILIA		TH3RESA	
	lapatinib plus capecitabine (n = 488)	T-DM1 (n = 490)	TPC (n = 198)	T-DM1 (n = 403)
Any Adverse Events	477 (97.7%)	470 (95.9%)	141 (76.6%)	337 (83.6%)
BLOOD AND LYMPHATIC SYSTEM DISORDERS				
Overall	87 (17.8%)	171 (34.9%)	54 (29.3%)	93 (23.1%)
THROMBOCYTOPENIA	12 (2.5%)	137 (28.0%)	6 (3.3%)	61 (15.1%)
NEUTROPENIA			40 (21.7%)	22 (5.5%)
ANAEMIA	39 (8.0%)	51 (10.4%)	19 (10.3%)	36 (8.9%)
GASTROINTESTINAL DISORDERS				
Overall	436 (89.3%)	352 (71.8%)	93 (50.5%)	224 (55.6%)
DIARRHOEA	389 (79.7%)	114 (23.3%)	40 (21.7%)	40 (9.9%)
NAUSEA	218 (44.7%)	192 (39.2%)	40 (21.7%)	133 (33.0%)
VOMITING	143 (29.3%)	93 (19.0%)	15 (8.2%)	71 (17.6%)
CONSTIPATION	47 (9.6%)	124 (25.3%)	29 (15.8%)	78 (19.4%)
DRY MOUTH	24 (4.9%)	77 (15.7%)	0 (0%)	49 (12.2%)
DYSPEPSIA	56 (11.5%)	43 (8.8%)		
ABDOMINAL PAIN UPPER	41 (8.4%)	57 (11.6%)	23 (12.5%)	26 (6.5%)
STOMATITIS	61 (12.5%)	16 (3.3%)	93 (50.5%)	224 (55.6%)
GENERAL DISORDERS AND ADMINISTRATION SITE CONDITIONS				
Overall	298 (61.1%)	331 (67.6%)	83 (45.1%)	198 (49.1%)
FATIGUE	136 (27.9%)	172 (35.1%)	46 (25.0%)	109 (27.0%)
ASTHENIA	81 (16.6%)	86 (17.6%)	29 (15.8%)	63 (15.6%)
MUCOSAL INFLAMMATION	93 (19.1%)	33 (6.7%)		
PYREXIA	37 (7.6%)	85 (17.3%)	22 (12.0%)	65 (16.1%)
INFECTIONS AND INFESTATIONS				
Overall	220 (45.1%)	213 (43.5%)		
PARONYCHIA	52 (10.7%)	1 (0.2%)		
INVESTIGATIONS				
Overall	139 (28.5%)	184 (37.6%)		
ASPARTATE AMINOTRANSFERASE INCREASED	46 (9.4%)	110 (22.4%)		
ALANINE AMINOTRANSFERASE INCREASED	43 (8.8%)	83 (16.9%)		
METABOLISM AND NUTRITION DISORDERS				
Overall	169 (34.6%)	144 (29.4%)	23 (12.5%)	58 (14.4%)
DECREASED APPETITE	113 (23.2%)	101 (20.6%)	23 (12.5%)	58 (14.4%)

System organ/ class/adverse events	EMILIA		TH3RESA	
	lapatinib plus capecitabine (n = 488)	T-DM1 (n = 490)	TPC (n = 198)	T-DM1 (n = 403)
MUSCULOSKELETAL AND CONNECTIVE TISSUE DISORDERS				
Overall	180 (36.9%)	249 (50.8%)	20 (10.9%)	79 (19.6%)
ARTHRALGIA	38 (7.8%)	85 (17.3%)	7 (3.8%)	46 (11.4%)
BACK PAIN	50 (10.2%)	64 (13.1%)		
PAIN IN EXTREMITY	52 (10.7%)	52 (10.6%)		
MYALGIA	18 (3.7%)	69 (14.1%)	15 (8.2%)	42 (10.4%)
NERVOUS SYSTEM DISORDERS				
Overall	189 (38.7%)	245 (50.0%)	15 (8.2%)	89 (22.1%)
HEADACHE	68 (13.9%)	133 (27.1%)	15 (8.2%)	89 (22.1%)
DIZZINESS	51 (10.5%)	48 (9.8%)		
NEUROPATHY PERIPHERAL	28 (5.7%)	49 (10.0%)		
PSYCHIATRIC DISORDERS				
Overall	74 (15.2%)	101 (20.6%)		
INSOMNIA	41 (8.4%)	54 (11.0%)		
RESPIRATORY, THORACIC AND MEDIASTINAL DISORDERS				
Overall	156 (32.0%)	217 (44.3%)	23 (12.5%)	97 (24.1%)
COUGH	60 (12.3%)	83 (16.9%)	19 (10.3%)	63 (15.6%)
EPISTAXIS	39 (8.0%)	99 (20.2%)	5 (2.7%)	47 (11.7%)
DYSPNOEA	36 (7.4%)	56 (11.4%)		
SKIN AND SUBCUTANEOUS TISSUE DISORDERS				
Overall	391 (80.1%)	159 (32.4%)	19 (10.3%)	19 (4.7%)
PALMAR-PLANTAR ERYTHRODYSAESTHESIA SYNDROME	283 (58.0%)	6 (1.2%)		
RASH	130 (26.6%)	52 (10.6%)	19 (10.3%)	19 (4.7%)
DRY SKIN	49 (10.0%)	17 (3.5%)		

Table 11 shows AEs of grade 3 or higher (AEs were graded according to the National Cancer Institute Common Toxicity Criteria for Adverse Events, version 3.0). There are slightly lower frequencies of AEs in the TH3RESA trial than in the EMILIA trial, probably reflecting a shorter time on treatment at time of analysis.

Table 11: AEs of grade 3 or higher, table adapted from MS

Adverse Event	% patients with grade 3 or higher event			
	EMILIA Lapatinib in combination with capecitabine n=488	EMILIA T-DM1 N=490	TH3RESA TPC n=184	TH3RESA T-DM1 n=403
Diarrhoea	20.7	1.6	4.3	0.7
Hand-foot syndrome	16.4	0		
Vomiting	4.5	0.8		
Neutropenia	4.3	2	15.8	2.5
Hypokalaemia	4.1	2.2		
Fatigue	3.5	2.4	2.2	2.0
Nausea	2.5	0.8		
Mucosal inflammation	2.3	0.2		
Thrombocytopenia	0.2	12.9	1.6	4.7
Increased AST	0.8	4.3	2.2	2.2
Increased ALT	1.4	2.9		
Anaemia	1.6	2.7	2.7	2.7
Abdominal pain			2.7	1.2
Asthenia			2.2	1.0
Cellulitis			2.2	0.5
Pulmonary embolism			2.2	0.5
Dyspnoea			1.6	2.0
Febrile neutropenia			3.8	0.2
Leukopenia			2.7	0.2

There were a number of fatalities while on study treatment, although percentages were considered low on both arms given the advanced cancer and associated ill health of the patients; EMILIA T-DM1 n=1 (0.2%) (metabolic encephalopathy); EMILIA lapatinib in combination with capecitabine n=4 (0.8%) (coronary artery disease, multi-organ failure, coma, hydrocephalus); TH3RESA T-DM1 n=5 (1.2%) (pneumonia, sepsis, hepatic encephalopathy, subarachnoid haemorrhage and pneumonitis); TH3RESA TPC n=3 (1.6%) (clostridium bacteremia, non-cardiogenic pulmonary oedema and pulmonary embolism).

In the EMILIA [4] trial, the T-DM1 group had fewer adverse events of grade 3 or greater, compared with those treated with lapatinib in combination with capecitabine: 40.8% (95% CI: 37% - 45%) versus 57% (95% CI: 53% - 61%). [4,6] A serious AE (SAE) was defined as any AE that resulted in death, was life-threatening, required in-patient hospitalization or prolongation of existing hospitalization, resulted in persistent or significant disability/incapacity, was medically significant, or required intervention to prevent any of these outcomes. [10] SAEs were experienced by 15.5% of the T-DM1 group, and 18.0% of the comparator group.

In the TH3RESA[5] trial, 32.3% of the T-DM1 group, and 43.5% of the TPC group, had adverse events of grade 3 or greater. Serious Adverse Events (SAEs) were experienced by 18.4% of the T-DM1 group, and 20.7% of the comparator group.

Trial JO22997,[9] shown within Table 9, had a population of Japanese patients with a median of 3 prior chemotherapy regimens for MBC (range, 1–8), including lapatinib in 43 (58.9%) patients. The abstract reports "The most frequently observed grade ≥3 adverse events were thrombocytopenia (21.9%), increased aspartate aminotransferase (13.7%), increased alanine aminotransferase (8.2%) and vomiting (5.5%).[9] One patient (1.4%) discontinued treatment due to thrombocytopenia. No patient received platelet transfusion. Grade 3/4 haemorrhage was observed in one patient (1.4%)".

In the pooled analysis submitted to the EPAR based on 882 patients, that was published as an abstract, the most common AEs were fatigue, nausea, headache, thrombocytopenia, and constipation.[10] The most common AEs of grade 3 or greater were thrombocytopenia, fatigue, increased hepatic transaminases (aspartate aminotransferase and alanine aminotransferase), hypokalemia, and anaemia.

The MS has provided an updated pooled analysis with 884 patients (cut-off date for pooled analysis 31.07.2012). Tables 12 and 13 are taken and adapted from Section 6.9 of the MS and the manufacturer's clarification response. See Appendix 4 of the ERG report for the Table of Common AEs in the pooled analysis of 884 patients treated with single-agent T-DM1, reproduced from the MS clarification response.

Table 12: Overview of serious AEs, adapted from the manufacturer's clarification response

Event	**n(%) from 884 patients given T-DM1 as single agent**
AEs leading to death	12 (1.4)
SAE	175 (19.8)
Grade ≥3 AE	398 (45.0)
AE leading to discontinuation of study drug	62 (7.0)

Table 13: Selected AEs in 884 patients given T-DM1 as single agent, adapted from MS clarification question response

Event	**n(%) AE**	**n(%) Grade ≥3 AE**	**n(%) SAE**	**AE leading to discontinuation of study drug**
Thrombocytopenia	285 (32.2)	105 (11.9)	8 (0.9)	15 (1.7)
Haemorrhage	323 (36.5)	18 (2.0)	14 (1.6)	-
Hepatotoxicity	92 (33.2)	81 (9.2)	10 (1.1)	18 (2.0)
Peripheral neuropathy	257 (29.1)	22 (2.5)	2 (0.2)	4 (0.5)
Infusion reactions/hypersensitivity	61 (6.9)	1 (0.1)	2 (0.2)	1 (0.1)
Cardiac dysfunction	14 (1.6)	2 (0.2)	-	2 (0.2)
Pneumonitis/Interstitial lung disease	10 (1.1)	3 (0.3)	3 (0.3)	4 (0.5)

4.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

Studies were selected according to the inclusion and exclusion criteria described in Section 4.1.6. Studies were excluded during sifting if the patients received prior therapy in the neo-/adjuvant setting rather than in the metastatic setting. One trial was excluded due to being a dose-ranging study (EGF10004).[11] Of trials deemed to meet the inclusion criteria, one RCT (EGF104900)[12] and two non-RCTs (Jerusalem 2011, a study of everolimus 5mg in combination with trastuzumab and vinorelbine versus everolimus 20mg in combination with trastuzumab and vinorelbine versus everolimus 30mg in combination with trastuzumab and vinorelbine; and Andre 2010, a study of everolimus 5mg in combination with paclitaxel and trastuzumab versus everolimus 10mg in combination with paclitaxel and trastuzumab versus everolimus 30mg in combination with paclitaxel and trastuzumab) were excluded for not linking into the network.[13,14]

Five RCTs were included in the MTC.[4,15-18] Trial characteristics of the RCTs included in the MTC are shown in Table 14 (adapted from the MS).

Table 14: Trial characteristics of the trials included in the MTC (Table adapted from the MS)

Study	**EMILIA trial**	**GBG26 Trial**	**EGF100151 trial**	**Martin *et al.* 2011**	**CEREBEL trial subgroup pre-treated with trastuzumab**
Primary study reference	Verma, 2012	Von Minckwitz, 2011	Cameron, 2008	Martin *et al.*, 2011	Pivot, 2012
Publication type	Journal article	Journal article	Journal article	Conference proceeding	Conference proceeding+ DHCP letter
Intervention	Capecitabine + Lapatinib (N=496)	Capecitabine + Trastuzumab (N= 78)	Capecitabine + Lapatinib (N=198)	Capecitabine + Lapatinib (N=116)	Capecitabine + Lapatinib (N=167)
Comparator (all active controlled)	T-DM1 (N=495)	Capecitabine (N= 78)	Capecitabine (N=201)	Neratinib (N=117)	Capecitabine + Trastuzumab (N= 159)
Location	USA and non-USA sites	Non-USA sites	USA and non-USA sites	NR	Non-USA sites
Prior therapy eligibility criterion	previously treated with trastuzumab and a taxane	progressed during treatment with trastuzumab with or without 1st-line metastatic chemotherapy, not required to have had taxane [n=42 no prior taxanes]	progressed after treatment with regimens that included, but were not limited to, an anthracycline, a taxane, and trastuzumab	required to have 2 prior trastuzumab regimens, prior taxane treatment	required to have received either a taxane *or* an anthracycline in the adjuvant setting; [unclear how many in subgroup not had prior taxanes] stratified based on their prior trastuzumab exposure (only those exposed to trastuzumab considered in MTC)
Design	RCT Phase III	RCT Phase III	RCT Phase III	RCT, Phase II	RCT Phase III
Method of randomisation	Adequate	Unclear	Unclear	Unclear	Unclear
Method of blinding (care provider, patient, outcome assessor)	Open-label but assessor-blind (IRC) for PFS	Open-label	Open-label, but assessor-blind (IRC) for TTP	Open-label	Open-label
Cross-over	No	No	Yes	No	No

Study	EMILIA trial	GBG26 Trial	EGF100151 trial	Martin *et al.* 2011	CEREBEL trial subgroup pre-treated with trastuzumab
permitted					
Primary outcome	PFS by IRC, OS, safety	TTP	TTP	PFS	CNS metastases
Secondary outcomes	PFS by INV, ORR, time to treatment failure, pharmacokinetics, DOR, patient-reported QoL, OS rate, TTP	OS, response rate, clinical benefit rate, DOR, safety, dose interruptions, withdrawal	PFS, OS, clinical benefit rate, withdrawal, safety, response rate, biomarker analysis	OS, safety, response rate, withdrawal, clinical benefit rate	PFS by INV, OS, ORR, CBR, time to first CNS progression, incidence of CNS progression at any time, safety
Present line of therapy: First-line, n (%)	0 (0)	NR	88 (22)	All patients were previously treated in the first or second-line setting	NR
Present line of therapy: first-line fast relapser, n (%)	118 (12)	NR	0 (0)		
Present line of therapy: Second-line, n (%)	361 (36)	156 (100)	NR		
Present line of therapy: third-line, n (%)	512 (52)	NR	NR	NR	0 (0)
Advanced or metastatic sites in the brain, n (%)	50 (10)	45 (9)	3 (2)	NR	NR

Study	**EMILIA trial**		**GBG26 Trial**		**EGF100151 trial**		**Martin *et al.* 2011**		**CEREBEL trial subgroup pre-treated with trastuzumab**	
Treatment group	Capecitabine + Lapatinib	T-DM1	Capecitabine+ Trastuzumab	Capecitabine	Capecitabine + Lapatinib	Capecitabine	Capecitabine + Lapatinib	Neratinib	Capecitabine + Lapatinib	Capecitabine+ Trastuzumab
Patients with ER+ and/or PR+, n (%)	155 (31)	176 (36)	41 (56) (N=73)	43 (62) (N=71)	96 (48)	93 (46)	NR	NR	NR	NR
Patients with Performance Status=1, n (%)	176 (36)	194 (39)	NR	NR	76 (38)	83 (41)	NR	NR	NR	NR
Study duration	Capecitabine + Lapatinib: 53.73 weeks (range: 0 weeks -151.67 weeks); T-DM1: 55.9 weeks (0 weeks - 147.33 weeks) – at first interim analysis Capecitabine + Lapatinib: 80.60 weeks (range: 0 weeks -177.67 weeks); T-DM1: 82.76 weeks (0 weeks - 173.33 weeks) – at second interim analysis[2]		89.70 weeks (20.7 months)		Capecitabine + Lapatinib: ~20 weeks; Capecitabine: ~15 weeks[2]		NR		NR	

The CEREBEL[18] and Martin[17] trials were excluded from the MTC within a sensitivity analysis due to differences from the other three trials in terms of patient characteristics (see section 4.4).

As shown within Table 14, there is variation between trials regarding patients' prior therapy. In GBG26[15] and CEREBEL not all patients had prior taxane treatment. The ERG's clinical experts believe that prior therapy with trastuzumab and/ or a taxane could be a clinically significant variable and thus could be a potential treatment effect modifier in the network meta-analysis. However, there are not enough trials to perform a meta-regression which would adjust for prior therapy.

There were some other differences between studies (apart from intervention and comparators):
Blinding: EMILIA [4] blinded outcome assessors for PFS, EGF100151[16] blinded outcome assessors for TTP, the other trials were not blinded for PFS/TTP.
PFS endpoint: Response evaluation criteria in solid tumours (RECIST) was used in EMILIA, Martin *et al*., and EGF100151 trials; it was unclear from CEREBEL and GBG26 whether this was used.
Despite comparability across trials in the number of patients with three or more metastatic sites, study populations differed in the sites of metastases.

4.4 Critique of the indirect comparison and/or multiple treatment comparison

The manufacturer conducted separate network meta-analyses for OS and PFS. The network of treatments included: lapatinib in combination with capecitabine; capecitabine monotherapy; trastuzumab in combination with capecitabine; and T-DM1, with the addition of neratinib when analysing PFS. Separate networks included data from four and five RCTs in the base case for OS and PFS respectively, which are shown in Figures 1 and 2.

Figure 1: Network of evidence for OS

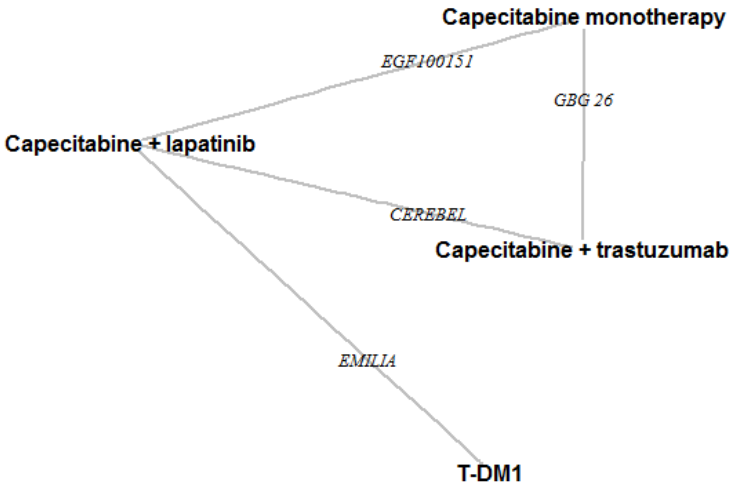
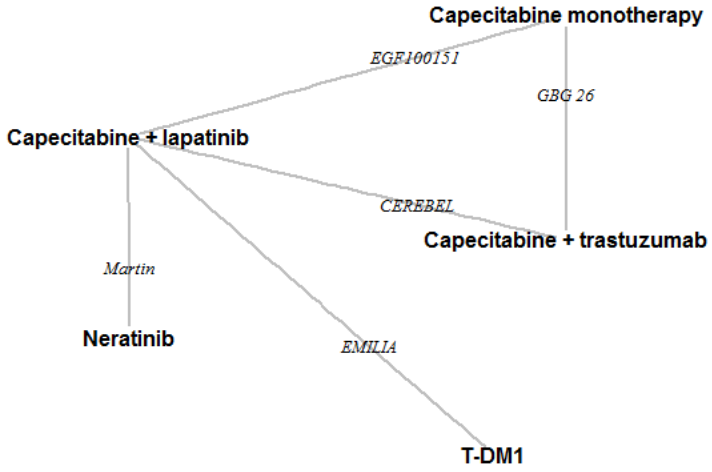


Figure 2: Network of evidence for PFS



The outcome measures of interest were the log hazard ratio for OS and PFS. The MS presented results from fixed effect models as follows:

- Estimated HRs and corresponding 95% credible intervals (CrIs) from network meta-analyses for all pairwise treatment comparisons for OS and PFS for:
    - The base case - including all identified trials for which data was available at the time of the literature search
    - A sensitivity analysis - excluding CEREBEL[18] (also excluding Martin *et al.*[17] for PFS)
- The probability of each treatment ranking
- Adjusted indirect treatment comparisons (ITC) (based on methods developed by Bucher *et al.*[19]), excluding CEREBEL

The log-hazard ratios and corresponding standard errors (SE) were synthesised in network meta-analysis using a Bayesian approach. The log-hazard ratios were then transformed to estimate the hazard ratios. The ERG agrees that synthesising the data in a network meta-analysis is appropriate because it quantifies the uncertainty in the parameters. The use of hazard ratios assumes that the treatment effect is constant over the lifetime of the patients. Based upon the analysis from the EMILIA trial within the MS, there is uncertainty around whether the assumption of proportional hazards holds for T-DM1 and lapatinib in combination with capecitabine beyond 72 weeks (see Section 5.2.6).

The base case results as presented in the MS (Figure 20, page 127) for OS is shown in Figure 3. For PFS, the fixed effect model results were recreated by the ERG because comparisons with neratinib were not presented in the MS, including comparisons with: trastuzumab in combination with capecitabine; trastuzumab emtansine; and capecitabine monotherapy. The ERG results for PFS are shown in Figure 4. The results of the network meta-analysis presented by the manufacturer suggested that T-DM1 is associated with a reduced hazard of both death and progression when compared with all other treatments in the comparator set.

Figure 3: Fixed effect model results for OS

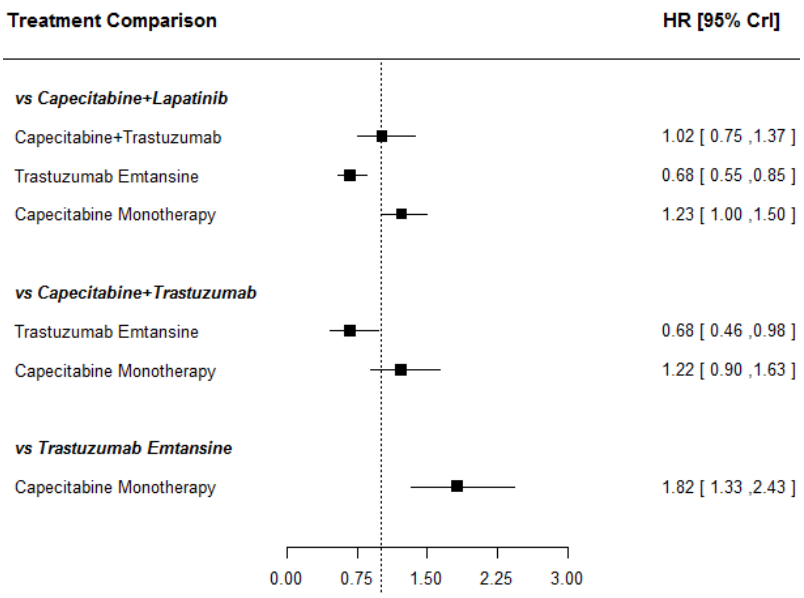
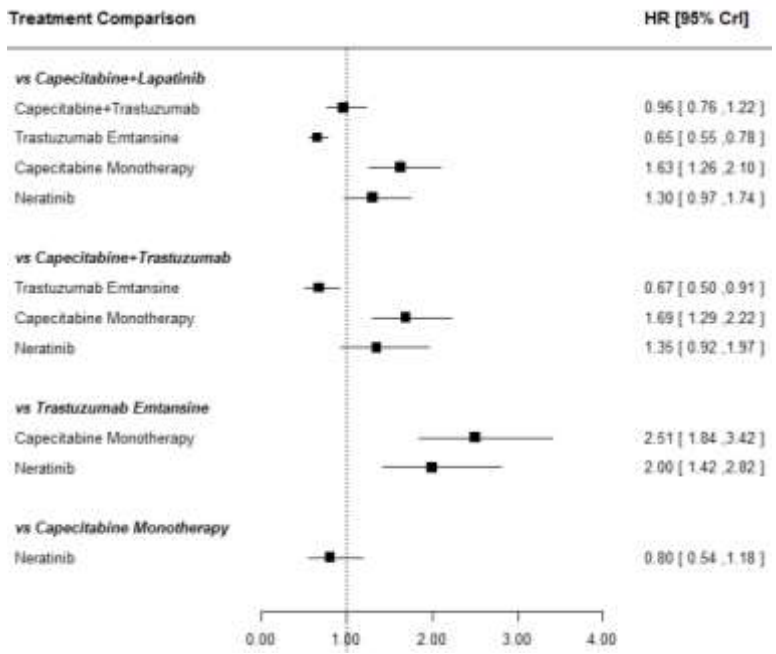


Figure 4: Fixed effect model results for PFS[†]



[†] *Recreated by ERG to include comparisons with neratinib*

The MS states that a fixed effect model was used to model the data rather than a random effects model because of the limited number of trials. However, this assumes that there is no heterogeneity between trials which is unlikely to be the case given the knowledge of the trials in the analysis. Within the clarification process, the ERG requested that the manufacturer present results using a random effects model using a weakly informative prior distribution for the between-trial standard deviation. In response to this, the manufacturer provided pairwise HRs from a random effects model but did not specify the prior distributions that were used. A standard reference prior distribution for the between-study standard deviation is Uniform(0,2), suggested by NICE DSU Technical Support Document 2 (Evidence Synthesis series).[20] However, this implies that extreme heterogeneity is equally plausible to mild heterogeneity, and will produce meaningless estimates of treatment effect in the absence of sufficient sample data to update the prior distribution. The ERG proposed using a weakly-informative prior distribution (i.e. a Half-Normal(0, 0.32^2)) for the between-study standard deviation as suggested in the NICE DSU Technical Support Document 3 (Evidence Synthesis Series).[21] The results from a model with this prior information are shown in Section 4.5.

The MS states that probabilities of treatment rankings were computed. However, this is only provided as raw WinBUGS output in Appendix 13 (Section 10.12.8 - Section 10.12.11) of the MS, with no interpretation of the findings. No probabilities of treatment rankings are provided by the manufacturer from the random effects model presented within the response to clarifications. The ERG provides these for OS and PFS in Section 4.5. Probabilities of each treatment ranking from the ERG random

effects model are different to the manufacturer's fixed effect results because probabilities depend on the whole posterior distribution. The manufacturer's ranking probabilities are presented in Figures 5 and 6 as shown numerically in Appendix 13 (Section 10.12.8 - Section 10.12.9) of the MS. T-DM1 has a 98% chance of being the best treatment for reduced hazard of death and 99% chance of being the best treatment for a reduced hazard of progression.

Figure 5: Probability of treatment rankings for OS

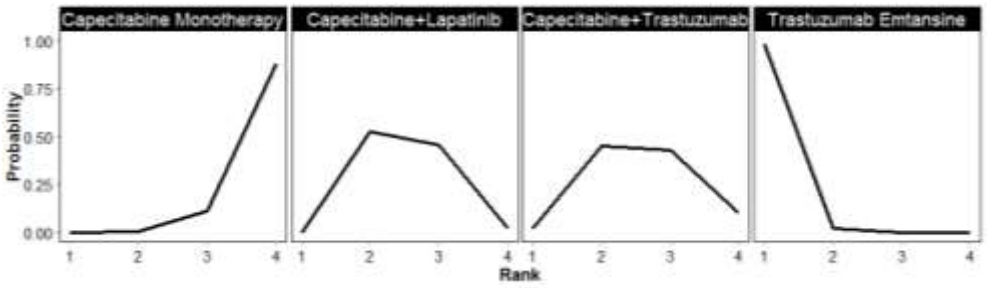
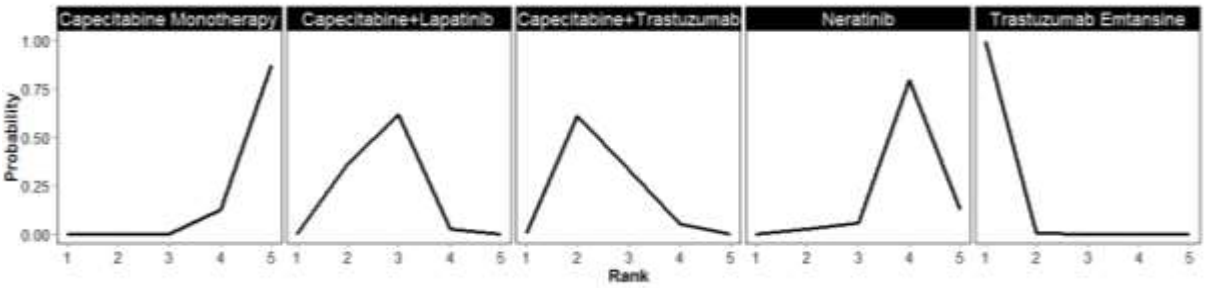


Figure 6: Probability of treatment rankings for PFS



There was one feedback loop involving lapatinib in combination with capecitabine, capecitabine monotherapy, and trastuzumab in combination with capecitabine in which it would be possible to assess inconsistency between direct and indirect estimates of treatment effect. Inconsistency would arise if there is imbalance in treatment effect modifiers comparing different pairs of treatments. The manufacturer did not assess inconsistency or discuss whether there was an imbalance in treatment effect modifiers.

Within the MS it is stated that after the initial literature review (December 2012), evidence was published from two trials, TH3RESA and BOLERO-3, relevant to the network. The TH3RESA trial is described in detail in Section 4.2. There is a disparity regarding the reason for exclusion of the TH3RESA trial (which would allow a comparison of T-DM1 with trastuzumab in combination with vinorelbine) in the MS and the clarification response. When asked by the ERG to clarify further why the larger network presented in Figure 18 of the MS was not used, the manufacturer stated that patients were randomised to either T-DM1 or TPC and the selection of therapy within the TPC arm

was made after randomisation. This means that there is no record of what therapy the patients randomised to T-DM1 would have received had they been randomised to the comparator arm. As the choice of therapy is highly influenced by a patient's characteristics (particularly characteristics indicative of their prognosis) it is not possible to make an unbiased, randomised comparison of T-DM1 and trastuzumab in combination with vinorelbine using this study. Comparing the two arms equates to a comparison of the ITT population of those randomised to T-DM1 to those selected to receive trastuzumab in combination with vinorelbine in the comparator arm. Whilst this issue could have been avoided by having the clinicians pre-specify the choice of alternative therapy (thereby allowing a comparison of those who would have received trastuzumab in combination with vinorelbine in both arms) this is unfortunately not the case. Given this reason and the fact that some of the treatment options within TPC are not listed as comparators within the NICE scope, the ERG believes that it is reasonable to exclude the TH3RESA trial from the MTC analysis.

The BOLERO-3 trial compared trastuzumab in combination with vinorelbine versus everolimus in combination with trastuzumab and vinorelbine. Since the initial literature review by the manufacturer was undertaken, data from the trial has been published but was not included in the analysis presented in the MS. However, data from BOLERO-3 would not have been synthesised in the analysis as the TH3RESA trial was also excluded and so the treatments in the BOLERO-3 trial would not have been connected in the network.

A sensitivity analysis excluding CEREBEL and Martin *et al*. was presented within the MS. The rationale for this analysis given in the MS was that "the heterogeneity assessment of these studies indicated that the patient population, prior treatment status and lack of detailed information on the study population's baseline characteristics in CEREBEL and Martin *et al*. deemed these two studies not entirely comparable to the other trials". The results of this analysis also suggest that T-DM1 is associated with a reduced hazard of both death and progression when compared with all other treatments in the comparator set. However, there is an inconsistency in the manufacturer's definition of the base case analysis between the MS and the clarification response. In the MS, the base case includes all trials and the sensitivity analysis excludes CEREBEL and Martin *et al.* However, in the clarification response, the base case excludes these two trials. The updated economic analysis given in Tables 5 and 6 of the clarification response uses the results after excluding the CEREBEL and Martin trials instead of the base case defined in the MS. The ERG believes that the base case should include the CEREBEL and Martin trials as defined in the MS, but should employ a random effects model to account for any between-study variability.

4.5 Additional work on clinical effectiveness undertaken by the ERG

The ERG repeated the MTC using a random effects model with a Half-Normal(0,0.32^2) prior distribution for the between-study standard deviation. The results of this analysis for all pairwise treatment comparisons are shown in Figures 7 and 8 for OS and PFS respectively. As expected for estimates of hazard ratios based on medians of posterior distributions, the pairwise HRs using the random-effects model are similar to those obtained using the fixed effect model (Figures 3 and 4). However, when the random effects model is used there is greater uncertainty induced by the between-trial variability and the possibility that there is no difference between treatments cannot be ruled out. The posterior estimates of the between-study standard deviation for OS was 0.18 (95% CrI: [0.01, 0.63]), and 0.18 (95% CrI: [0.01, 0.62]) for PFS, which is indicative of mild heterogeneity.

Figure 7: Random effects model for OS

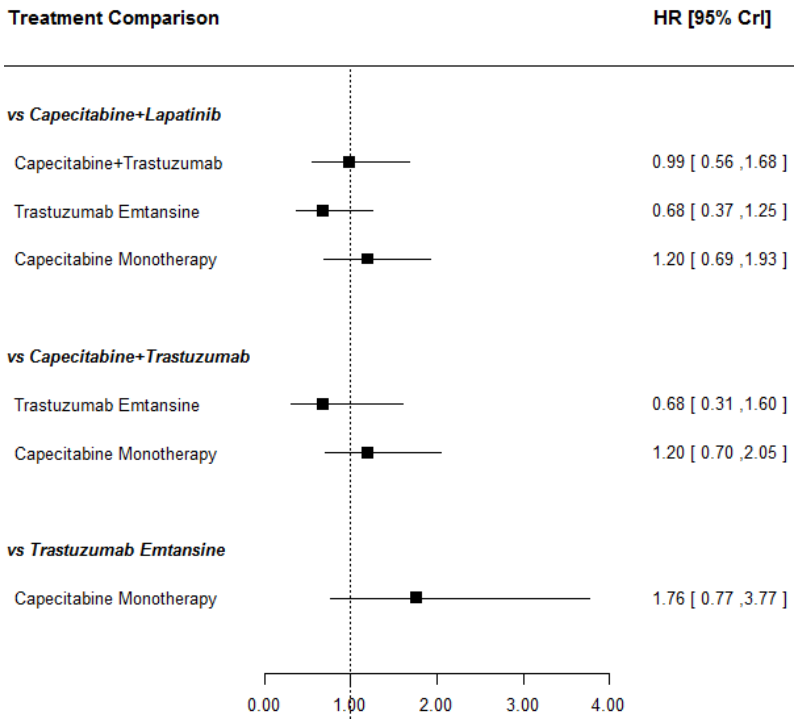
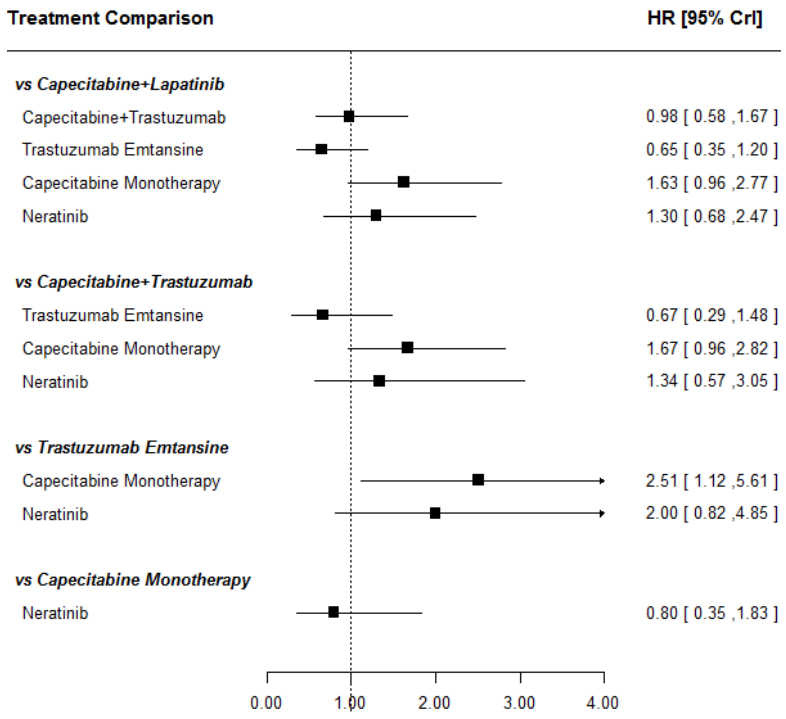


Figure 8: Random effects model for PFS



The probabilities of each treatment ranking for OS and PFS are shown in Figures 9 and 10. Using a random effects model, the probabilities of trastuzumab being the best treatment in terms of the hazard for OS and PFS are 84% and 87% respectively.

Figure 9: Probability of treatment rankings for OS

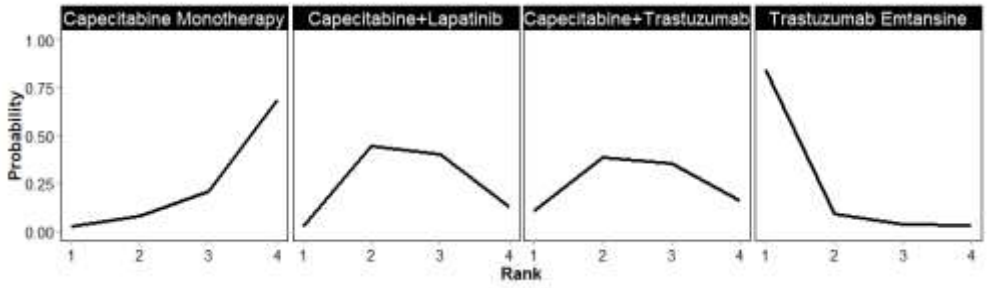
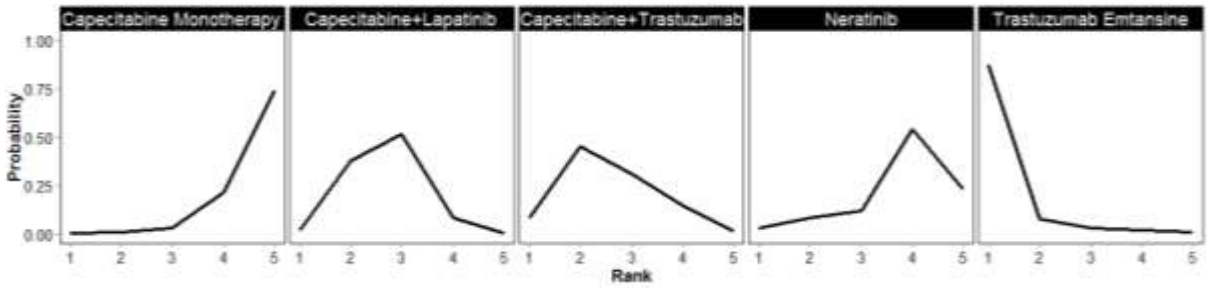


Figure 10: Probability of treatment rankings for PFS



4.6 Conclusions of the clinical effectiveness section

A systematic review was conducted for the decision problem, identifying two relevant trials. The ERG believes that all relevant trials with available data have been included. The clinical effectiveness data relevant to the decision problem were taken from two large RCTs, both of which were open-label, but otherwise at low risk of bias. The lack of blinding is unlikely to have affected OS, but could bias the HRQoL data. One of the two trials had independent outcome assessment of PFS. Although both trials were ongoing at the time of writing, they had completed their primary endpoint. Data were available for OS, PFS and AEs, and one trial provided HRQoL data (although not in a form that can be transformed to utility values). Additionally, adverse event data were available from a pooled analysis of T-DM1 trials.

Data from the two RCTs of T-DM1 reported a significant advantage in PFS for T-DM1 over lapatinib in combination with capecitabine, and over TPC. Data from one RCT reported a significant advantage in OS for T-DM1 over lapatinib in combination with capecitabine. Further OS analyses of both trials are planned (at the time of this assessment). There was some suggestion of improvement in time to symptom worsening for T-DM1 over lapatinib in combination with capecitabine.

The most common AEs for T-DM1 were fatigue, nausea, headache, thrombocytopenia, and constipation. The most common AEs of grade 3 or greater were thrombocytopenia, fatigue, increased hepatic transaminases (aspartate aminotransferase and alanine aminotransferase), hypokalaemia, and anaemia. Adverse event data from RCTs showed fewer AEs of grade 3 or greater for T-DM1 than for lapatinib in combination with capecitabine, or than for treatment of physician's choice. Frequencies of SAEs were broadly similar, although slightly lower in T-DM1 groups. Limited HRQoL was available, although FACT-B TOI-PFB data collected within the EMILIA trial suggested deterioration took longer in the T-DM1 group than the comparator.

Most of the data were from third-line or later therapy, whereas the MS suggests T-DM1 as second-line treatment. The studies were international, so not all participants would have had prior treatment in accordance with UK practice. The trial populations were broadly similar to populations that would be encountered in UK practice, although in practice there may be more patients with ECOG PS2.

The only comparator from the final NICE scope for which there was head-to-head data, was lapatinib in combination with capecitabine. There was a lack of head-to-head comparison with T-DM1 for all other comparators in the decision problem. Within the MTC analysis, allowing for heterogeneity between studies increases the uncertainty about the true treatment effect on OS and PFS. T-DM1 appears to be the best treatment in terms of both OS and PFS. From the ERG's random effects model, T-DM1 is associated with a reduction in the hazard of death of 32% (HR=0.68, 95% CrI [0.37, 1.25])

and a reduction in the hazard of progression or death of 35% (HR=0.65, 95% CrI [0.35, 1.20]) compared to lapatinib in combination with capecitabine.

5. COST EFFECTIVENESS

5.1 ERG comment on manufacturer's review of cost-effectiveness evidence

A search was conducted by the manufacturer to identify published cost effectiveness evidence and to determine if any new modelling was required. The search was developed by using key references identified through scoping searches. Free text terms reflecting the population and intervention concepts of the decision problem were combined with economic evaluation / model outcome terms. Medline, Embase, Embase Alert and NHS-EED were the data sources for this search. EconLit was not searched, the manufacturer have acknowledged through the clarification process that this was an oversight. The search was limited from 1993 to current (which was 04.10.13). No supplementary techniques were reported such as citation or reference searching. In addition, no subject headings were used, which could limit the sensitivity of the search. The manufacturer acknowledged through the clarification process that this was an oversight. A verifiable cost-effectiveness study filter was not used, but terms designed to retrieve the appropriate study type were employed. No relevant studies were identified.

It is unlikely that any economic evaluations of T-DM1 have been missed by the manufacturer. However, in order to find other potentially useful evidence, a search statement reflecting comparator terms could have been added to the search to find studies on either T-DM1 or any of the comparators. With a lack of economic evidence found on T-DM1 it may have been possible to use economic evaluations of the comparator drugs to inform the manufacturer's model. Due to time constraints it was not possible for the ERG to devise a new search strategy incorporating comparator terms but a preliminary search adding comparator terms and relevant subject headings retrieved a manageable 58 references in Medline / Medline in Process, some of which could have been used to inform the economic model. Ideally the manufacturer's search would be expanded before reaching a decision on the status of the evidence base.

5.2 Summary and critique of manufacturer's submitted economic evaluation by the ERG

The manufacturer submitted a cohort state transition model written within Microsoft Excel ®. The main model structure is clinically appropriate and the implemented model is generally clear, with no major errors identified.

5.2.1 NICE reference case checklist

The manufacturer's economic evaluation follows the NICE Reference Case, as shown within Table 15 below, taken from the MS.

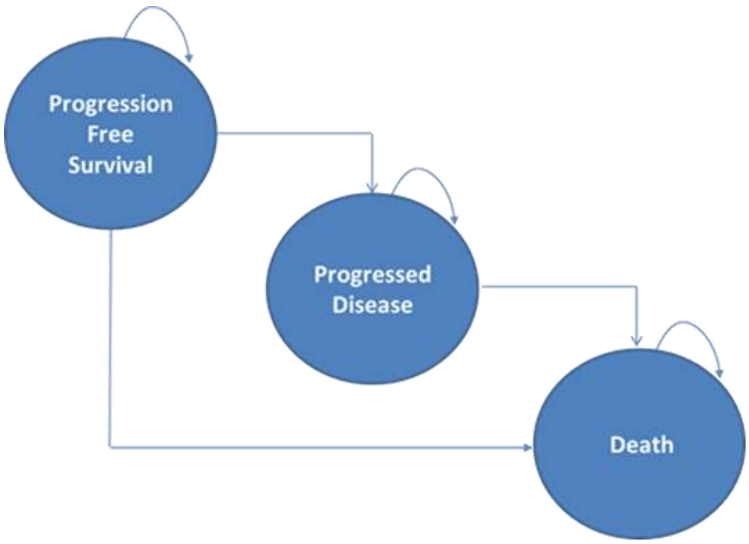
Table 15: Comparison of the MS with the NICE Reference Case checklist

Element of health technology assessment	**Reference Case**	**Does the submission adequately address the Reference Case?**
Defining the decision problem	The scope developed by the Institute	Yes
Comparator	Therapies routinely used in the NHS, including technologies regarded as current best practice	Yes
Perspective on costs	NHS and PSS	Yes
Perspective on outcomes	All health effects on individuals	Yes
Type of economic evaluation	Cost-effectiveness analysis	Yes
Synthesis of evidence on outcomes	Based on a systematic review	Yes
Measure of health effects	QALYs	Yes
Source of data for measurement of HRQoL	Reported directly by patients and/or carers	Yes
Source of preference data for valuation of changes in HRQoL	Representative sample of the public	Yes
Discount rate	An annual rate of 3.5% on both costs and health effects	Yes
Equity weighting	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes

5.2.2 Model structure

The cohort state transition model has three health states: PFS; progressed disease; and death. A clinically appropriate cycle length of one week has been employed. Each cycle, patients can either transition from PFS to progressed disease, remain in the current state, or transition to death, as shown within Figure 11 (replicated from the MS).

Figure 11: Model Structure



5.2.3 Population

The population within the model is adult patients with HER2-positive, unresectable LABC or MBC who previously received trastuzumab and a taxane, separately or in combination. This is consistent with the final scope. The patients simulated in the model were assumed to have a mean age of 53 years.

5.2.4 Interventions and comparators

The intervention, T-DM1, is compared with: lapatinib in combination with capecitabine; capecitabine; vinorelobine; trastuzumab in combination with capecitabine; and trastuzumab in combination with vinorelbine, as described within the NICE scope.

5.2.5 Perspective, time horizon and discounting

The model takes a NHS and PSS perspective as per the NICE reference case. Patients are followed over 10 years within the MS base case. However, this was increased to 15 years within the clarification process since up to 3% of patients remain alive at 10 years. By 15 years more than 99% of patients have died. Costs and outcomes are discounted at 3.5% per annum.

5.2.6 Treatment effectiveness and extrapolation

The effectiveness of T-DM1 and lapatinib in combination with capecitabine is based upon the EMILIA trial comparing these two treatment options. The data cut points were from January 2012 and July 2012 for PFS and OS respectively, which is prior to treatment switching. The MS assesses a wide range of options for extrapolating the PFS and OS data, including the use of parametric distributions and direct use of the Kaplan-Meier estimates in combination with parametric distributions for the tails of the curves. Within the MS base case, for PFS the Kaplan-Meier curve is applied directly until week 72, after which a lognormal distribution is used to represent the tail of the curve, whilst for OS the gamma distribution is fitted for the entire curve. The decision about which extrapolation approach to use within the base case is based upon cumulative hazard plots, visual fit, external validity and clinical plausibility, as recommended by Latimer within a NICE DSU Technical Support Document.[22] In order to adjust the lognormal distribution (or other parametric distributions within the sensitivity analysis) so that it meets the Kaplan-Meier curve at week 72, the transition probabilities for each weekly cycle from the lognormal distribution are multiplied by the proportion of patients remaining in PFS from the previous cycle, rather than the absolute figures generated from the lognormal distribution being used.

For PFS, The Kaplan-Meier curve was used because the fit to the observed data of all of the parametric distributions tested was shown not to be good. However, the manufacturer justifies the use of a lognormal distribution to estimate the tail of the curve based upon this being the best fit to the observed data. A more appropriate criterion would be to use the most clinically valid distribution beyond the observed data.[22]

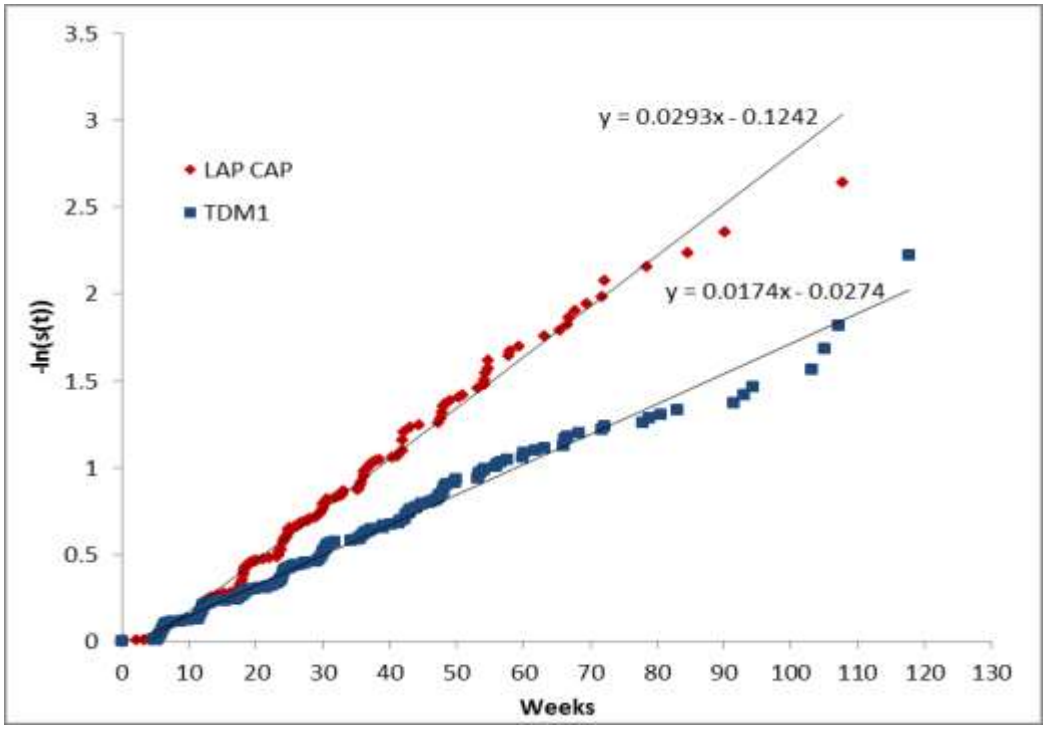
The effectiveness of capecitabine monotherapy and trastuzumab in combination with capecitabine is based upon the results of the MTC. This analysis provides HRs for the comparators compared with T-DM1. The use of a HR assumes that the treatment effect is constant over time so that hazards are proportional. The lognormal distribution is used within the base case for extrapolating the tail of the PFS data, although a lognormal distribution is an accelerated failure time model rather than a proportional hazards model. In combination with the concern raised above about the choice of the lognormal distribution for the tail of the curve, it may have been preferable to use an alternative distribution which is consistent with the proportional hazards assumption, such as the Weibull distribution. However, in practice the use of the Weibull distribution results in patients spending longer within the progressed disease state which is less clinically plausible (2.25 years for the lognormal versus 2.48 years for the Weibull for T-DM1 when the Gamma is used for OS). Thus the use of the lognormal distribution generates a clinically plausible curve even though the application of a hazard ratio to a lognormal distribution is theoretically incorrect.

Similarly for OS, the proportional hazards assumption does not hold for a Gamma distribution apart from the special case where it is equivalent to a Weibull distribution. However, the Gamma distribution provides a better fit to the OS data and has greater external validity beyond the trial data than any of the other distributions tested. Thus, whilst the assumptions are theoretically inconsistent, the use of the Gamma distribution for OS may give the most plausible results given the data available. The impact of the use of alternative distributions is tested within the ERG's sensitivity analysis (see Section 5.3 and Section 6).

Given the absence of relevant trials of vinorelbine identified for inclusion within the MTC, the HR for vinorelbine for PFS and OS is assumed to be the same as capecitabine. This assumption is based upon NICE Clinical Guidelines 81 where the guideline development group agreed that the effectiveness of the two treatments is essentially equivalent.[1] Similarly, the HR for PFS and OS for trastuzumab in combination with vinorelbine was assumed to be the same as trastuzumab in combination with capecitabine.

Importantly, the MS suggests that the comparative effectiveness between T-DM1 and lapatinib in combination with capecitabine may not be constant beyond week 72 where the points fit less well to the linear regression, as shown within Figure 12, replicated from the MS. This means that applying a HR for the comparators may result in inaccurate estimates of effectiveness over time. However, as the manufacturer suggests, this may be because of the low number of events occurring beyond 72 weeks. Given the limited data available for the comparators over time, the assumption of constant hazards provides a practical option for extrapolation; however the uncertainty around this should be noted.

Figure 12: Cumulative hazard plot for PFS for T-DM1 compared with lapatinib in combination with capecitabine



In addition, the model assumes that the treatment effect will be maintained over time, beyond the EMILIA trial data. This is subject to substantial uncertainty. The ERG has undertaken a sensitivity analysis testing the extreme assumption that there is no benefit of T-DM1 beyond the trial data (see Section 5.3 and Section 6).

Finally, the model assumes no relationship between PFS and OS. This means that estimates of PFS could be greater than estimates of OS within the stochastic model, although in practice this does not occur due to the substantial difference between PFS and OS and the limited uncertainty around PFS as implemented by the manufacturer.

5.2.7 Health related quality of life

No HRQoL data has been published from the trials which can be converted into utilities, although EQ-5D data is a secondary outcome of TH3RESA and interim results have been provided confidentially within the supplementary files of the MS. An update to an existing search for utility values from the NICE pertuzumab Single Technology Appraisal (STA)[17] was presented within the MS which did not identify any additional relevant studies. The HRQoL search did not use a verifiable study type filter. If such a filter were used in conjunction with terms for disease-specific instruments, such as the Functional Assessment of Cancer Therapy-Breast (FACT-B), it would have increased the

yield and possible sensitivity of the search. The sources were appropriate (Medline, Embase, Embase Alert and NHS EED), although a specialist data source such as the Tuft's medical centre CEA Registry would be a useful addition to this list. Terms for breast cancer and metastatic breast cancer could be combined with "or" to increase yield.

Due to the lack of evidence identified by the HRQoL search, the HRQoL within the model is estimated using a statistical model by Lloyd *et al*.[23] This mixed model analysis was based upon a sample of 100 people from the general population of England and Wales who were asked to value different health states and adverse events associated with metastatic breast cancer using the standard gamble technique.[23] Included variables are: age; treatment response; disease progression (according to whether patients are in the PFS or progressed disease state); febrile neutropenia; diarrhoea and vomiting; hand-foot syndrome; stomatitis; fatigue; and hair loss. Utilities are estimated using a treatment response variable of 0 (stable disease) and 1 (response), and subsequently weighted according to the objective response rates reported within the trials.

The utility values employed within the model based upon Lloyd *et al.* are shown within Table 16 below. These are consistent with the results of a systematic review and meta-analysis of health state utility values in metastatic breast cancer by Peasgood, Ward and Brazier (2010) which reports values between 0.721 and 0.806.[24] In addition, the ERG's clinical advisors agree that quality of life is likely to be greater for patients on T-DM1 because of the reduced adverse event profile.

Table 16: HRQoL employed within the MS

State	**Utility value – from MS**
PFS T-DM1	0.78
PFS lapatinib in combination with capecitabine	0.74
PFS trastuzumab in combination with capecitabine	0.73
PFS capecitabine	0.72
PFS trastuzumab in combination with vinorelbine	0.73
PFS vinorelbine	0.72
Progressed disease	0.50

Although these values have reasonable external validity, the ERG has a number of concerns with the way in which they are calculated, and thus they are subject to some uncertainty.

First, within the model, an age of 47 years has been used to calculate utilities and this is not dependent upon the age of patients within the model. The manufacturer states that age 47 is from Lloyd *et al.*; however the mean age reported within this publication is 40 years.[23] This mixed model analysis contains age as a variable so that this can be altered according to the patients of interest. However, the analysis by Lloyd *et al.* suggests that increasing age has a positive impact upon HRQoL. This contradicts other established sources based upon larger samples of the general population which suggest that HRQoL generally decreases with age.[25] Given this, in combination with the subsequent issues below and the reasonable external validity of the utilities employed, the age variable has not been altered by the ERG within the model.

Second, the adverse events included within the mixed model analysis by Lloyd *et al.* are not directly comparable with the serious adverse events experienced by patients on T-DM1 or its comparators. The MS states that quality of life impacts associated with grade 3 and 4 adverse events with over 2% incidence in either treatment arm of EMILIA are included within the model. According to Table 21 of the MS, this should include diarrhoea, hand-foot syndrome, vomiting, neutropenia, hypokalaemia, fatigue, nausea, mucosal inflammation, thrombocytopenia, increased AST, increased ALT and anaemia, many of which are not included within the model by Lloyd *et al*.[23] This means that the utility values do not account for many of the adverse events. Moreover, the frequency of each AE has been multiplied by the coefficients from the model by Lloyd *et al.* rather than using a binary variable, experience AE or not, and then weighting the total according to the frequency of that AE, as would be appropriate for a mixed model analysis.

A further limitation with the utility values is that the ERG's clinical experts suggest that the assumption of vinorelbine and capecitabine having the same adverse event profile is not clinically valid since these different treatment options lead to different types of adverse events. In addition, within their submission the manufacturer highlights that the utility associated with progressed disease may be an underestimate.

Given these issues, the ERG has undertaken a sensitivity analysis varying the utility values within plausible ranges to assess the impact of uncertainty within these parameters (see Section 5.3 and Section 6).

5.2.8 Resources and costs

The health economic model includes: the cost of the treatments; the cost of administration; the cost of treating a selection of AEs; supportive care costs; and costs of treatment within the progressed disease state. A systematic search was undertaken by the manufacturer for resource use data from the past five years from a UK NHS perspective for advanced or metastatic breast cancer. This was an update to an existing search described within the NICE pertuzumab STA,[26] designed to identify any newly published data since March 2013 until October 2013. No new resource use data was identified.

5.2.8.1 Cost of the treatments

The unit costs of TDM-1, trastuzumab, lapatinib, capecitabine and vinorelbine are based upon the latest BNF estimates,[27] which is an established source.

The dosage of lapatinib is consistent with the recommended dose within the BNF. The dosages of T-DM1, trastuzumab, capecitabine and vinorelbine are all based upon the weight or body surface area of the patient. For T-DM1, patients require a three-weekly dose of 3.6mg per kg. For trastuzumab, patients require 8mg per kg for the initial dose, followed by a three-weekly maintenance dose of 6mg per kg. For both of these treatments, the model assumes that vial sharing will not occur and therefore any excess treatment within the vials will be wasted. The average weight of a patient within EMILIA (70.1kg) is used to estimate these costs. When wastage is included, this does not consider the variation in weight between people and the impact of this upon vial usage. For example, for T-DM1, for individuals weighing 75kg rather than 70.1kg, two 160mg vials would be required rather than one 160mg and one 100mg vial, increasing the annual cost from £73,955 to £91,021. For T-DM1, because the average weight within the trial is closer to the threshold for increasing vial usage than that for decreasing it (i.e. in order to receive one 160mg vial, the woman would have to weigh less than 44.4kg), the total cost of the drug is underestimated. Thus, an alternative cost has been calculated by the ERG for both T-DM1 and trastuzumab based upon an approximated weight distribution of patients (see Section 5.3 and Section 6).

It should be noted that trastuzumab may now alternatively be administered subcutaneously as a fixed dose of 600mg every 21 days.[28] The impact of the reduced costs associated with this alternative form of administration is tested within the ERG's univariate sensitivity analysis (see Section 5.3 and Section 6).

Patients being treated with lapatinib in combination with capecitabine require 1000mg per m^2 of capecitabine twice daily for 14 out of 21 days, whilst those being treated with capecitabine monotherapy require 1250mg per m^2. Capecitabine is given to patients as 500mg or 150mg tablets.

An alternative cost has been calculated by the ERG for capecitabine based upon an approximated body surface area distribution of patients. For vinorelbine, the model assumes that vial sharing is commonplace because it is available to the NHS in generic form, thus it is not necessary to recalculate the cost of vinorelbine to adjust for weight variation.

The cost of treatments within the manufacturer's base case is based upon the planned dose rather than the actual dose of the drugs. Since the effectiveness estimates for T-DM1 and lapatinib in combination with capecitabine are based upon the actual dose provided and the actual dose of capecitabine was substantially lower than the planned dose within EMILIA, the ERG proposes using the actual dose provided within the base case where possible.

5.2.8.2 Cost of administration

The cost of administration includes the cost of administering the treatment and pharmacy costs, weighted according to whether the treatment is intravenously or orally administered. These costs are based upon PSSRU unit costs,[29] an established source, inflated to 2013 prices. Within the model, there is an error within the coding for the administration of trastuzumab in combination with vinorelbine. Trastuzumab is administered every three weeks whilst vinorelbine is administered on a weekly basis. The submitted model code multiplies the cost of administration of trastuzumab in combination with vinorelbine by the proportion of patients in PFS during that cycle and subtracts half of the pharmacy costs every three weeks, which results in incorrect costs of administration including some negative weekly costs. This is corrected within the ERG's base case (see Section 5.3 and Section 6).

5.2.8.3 Supportive care costs

A weekly cost is applied throughout the progression-free and progressed disease health states to account for a fortnightly community nurse visit (20 mins), a monthly GP visit and a monthly visit to a clinical nurse specialist (1 hour). Our clinical experts suggest that this is reasonable. These costs are based upon uplifted PSSRU unit costs.

The MS suggests that patients receiving HER2-directed therapies (TDM-1, trastuzumab and lapatinib) should have regular monitoring of left ventricular ejection fraction (LVEF). The cost of monitoring is not included within the health economic model. However, based upon NICE Clinical Guidelines 81 this would only be required every 3 months at a cost of around £130.[1] Thus this is unlikely to impact upon the model results substantially because it is a minimal cost relative to drug acquisition costs.

5.2.8.4 Costs of treatment within progressed disease state

Within the model, patients may receive capecitabine and/or vinorelbine within the progressed disease state, followed by palliative care. In theory, the patients are divided into whether they received TDM-

1/ its comparators as first-line, second-line or third-line treatment according to the proportions within EMILIA. After progression, first-line patients (12%) are assumed to receive 38 weeks of treatment (19 weeks on vinorelbine and 19 weeks on capecitabine); second-line patients (36%) are assumed to receive 19 weeks of treatment (half on capecitabine and half on vinorelbine); and those failing on third-line (52%) are assumed to receive no further active treatment following progression. Within the model, the same cost is applied for each week within the progressed disease health state. This value is calculated as the total cost of post-progression treatments divided by the mean time in the progressed disease state to spread out these costs over the remaining time within the model. This is simplified within the model because only one health state is employed to represent progressed disease and a more complex model structure would need to be employed to incorporate different treatment options over time. Given the expected impact of this cost upon the model results, this simplified approach seems reasonable were it to be modelled appropriately. However, the method used by the manufacturer assumes the weekly cost in the progressed disease state is independent of treatment. This results in those treatments, such as TDM-1, where patients spend a longer duration in the progressed disease state, being associated with greater costs than those with shorter durations, despite having similar post-progression treatments. The ERG has amended this error by calculating average costs per week for each individual treatment.

In addition, there is a lack of external validity associated with patients remaining in the progressed disease state for an average of 1.2 – 2.5 years (depending upon treatment within the PFS state) whilst only receiving active treatment for a maximum of 38 weeks. Thus the post-progression costs may be underestimated within the model. The cost associated with the progressed disease health state is increased by the ERG within a sensitivity analysis (see Section 5.3 and Section 6).

Palliative care costs are included within the model as a single cost upon death. This is estimated from the DIN-LINK database (an anonymised database of individual primary care records from general practices in the UK) using costs from 2000 – 2001, uplifted to 2013 prices. The ERG's clinical advisors suggest that there may now be more involvement by hospital and hospice teams than in 2000; however since this cost is applied to all patients within the model it will not impact substantially upon the model results.

5.2.8.5 Cost of adverse events

The cost of treatment of adverse events is based upon NHS Reference Costs, which is an established source; although all appear to be based upon 'Malignant Breast Disorders with Major CC (reduced short stay emergency tariff)' and other codes may be more appropriate for some AEs. The MS states that the cost of treatment associated with those grade 3 and 4 adverse events with over 2% incidence in either treatment arm of EMILIA are included. According to Table 21 of the submission, this

includes diarrhoea, hand-foot syndrome, vomiting, neutropenia, hypokalaemia, fatigue, nausea, mucosal inflammation, thrombocytopenia, increased AST, increased ALT and anaemia. However, within the model only costs associated with diarrhoea and fatigue are included. Costs associated with increased AST, hand-foot syndrome and thrombocytopenia are assumed to be zero, whilst the remaining adverse events are not mentioned in relation to the model. Since more grade 3 and 4 adverse events are experienced by those patients receiving lapatinib in combination with capecitabine than those receiving T-DM1, this is likely to be unfavourable to T-DM1.

The cost of AEs from the EMILIA trial is assumed to be spread equally over the remaining time within PFS. Given that all of these adverse events are currently experienced within the follow up period of EMILIA (median 19.1 months for T-DM1; median 18.6 months for lapatinib in combination with capecitabine), the costs associated with AEs are expected to be underestimated within the model because of the effect of discounting and because there are likely to be additional adverse events beyond the current trial follow up for patients staying in PFS. However, the sensitivity analysis around costs of AEs undertaken by the ERG suggests that this is unlikely to have a substantial impact upon the model results (see Section 5.3 and Section 6).

The weekly cost of AEs is calculated based upon summing the following for all types of AEs: (Total number of AE/Total time on treatment)*Cost of treating AE. It is unclear within the model from where the value for total time on treatment has been derived and the manufacturer did not clarify this satisfactorily when asked during the clarification process. Using the mean time on treatment within Sheet "KM TOTT" of the manufacturer's model and multiplying this by the number treated within Sheet "Demographic" provides similar figures to those included within the model, as shown in Table 17, although it is noted that the manufacturer's values are more favourable to lapatinib in combination with capecitabine.

Table 17: Total time on treatment (all patients, weeks)

Treatment	Manufacturer's figures	ERG figures
T-DM1	18,144	19,110
Lapatinib in combination with capecitabine	15,377	13,943

Within the economic model, the calculations for total weekly AE costs are multiplied by the proportion of patients on treatment rather than the proportion of patients in the PFS health state. This does not seem to be appropriate because the ERG believes that the weekly cost of adverse events has already taken into account time on treatment until progression. Thus within the ERG's base case analysis, the weekly costs of AEs are multiplied by the proportion of patients in PFS rather than the proportion on treatment.

5.2.9 Cost effectiveness results

In the initial submission, the manufacturer did not present a full incremental analysis, although this was corrected in the response to clarifications. These results are replicated in Table 18. Only deterministic results were presented within the clarification response. These are similar to the probabilistic results re-run by the ERG; however, the uncertainty around the model inputs was inadequately characterised (see Section 5.2.10).

Table 18: Replicated deterministic revised incremental cost-effectiveness analysis results from manufacturer's clarifications

Technologies	Totals			Incrementals			
	Costs (£)	LYs	QALYs	Costs (£)	LYs	QALYs	ICER (Cost per QALY gained)
Capecitabine	£11,850	1.61	0.89				
Vinorelbine	£16,518	1.61	0.89	£4,668	0.00	0.00	Dominated
Lapatinib and capecitabine	£34,227	2.53	1.45	£17,709	0.92	0.57	£39,449
Trastuzumab and capecitabine	£35,784	2.24	1.28	£1,557	-0.29	-0.17	Dominated
Trastuzumab and vinorelbine	£36,662	2.24	1.28	£878	-0.29	-0.17	Dominated
T-DM1	£111,226	3.16	1.91	£74,565	0.92	0.63	£167,253

It should be noted that the mean hazard ratio from the MTC for lapatinib in combination with capecitabine versus trastuzumab in combination with capecitabine is close to 1 for both PFS and OS, suggesting similar efficacy between these treatment options. However, estimating PFS and OS based upon the Kaplan-Meier data from EMILIA, the model predicts that lapatinib in combination with capecitabine is substantially more efficacious. Since the cost of trastuzumab in combination with capecitabine is greater than that for lapatinib in combination with capecitabine, even if PFS and OS were the same for these two treatment options, trastuzumab in combination with capecitabine would remain dominated.

5.2.10 Sensitivity analyses

The manufacturer undertook PSA and several univariate sensitivity analyses. However, both analyses have flaws.

Within the MS, the parameters of the distributions representing the uncertainty are not presented or justified and not all parameters varied within the PSA are described. Only costs and utilities are listed as being included within the PSA, yet within the economic model the parameters associated with the parametric distributions for extrapolation of PFS and OS and the hazard ratios for the comparators are also included within the PSA. Based upon the parameters varied within the economic model, the characterisation of uncertainty generally appears to be arbitrary.

Within the clarification process, the manufacturer was asked to provide the Excel spreadsheet referenced within the economic model for the parameters used to produce the parametric distributions for modelling PFS and OS, in order to clarify how the parameters and correlation matrices pasted as values within the model were derived. The manufacturer did not supply this; however based upon the data within the model, the ERG believes that the regression approach used to estimate the parameters for PFS and OS produces correlation matrices which are arbitrary rather than appropriately characterising the uncertainty. In addition, the MS acknowledges that no relationship is assumed between PFS and OS, which in theory means that estimates of PFS could be greater than estimates of OS within the PSA.

The uncertainty around the HRs does not take into account the joint distribution of treatment effects as generated by the MTC. The manufacturer approximates the posterior distribution of treatment effects using normal distributions with means and standard errors of log-hazard ratios for each treatment derived from the network meta-analysis and ignores the correlation between treatment effects. This is an unnecessary assumption that could be managed by including a look-up table of samples from the joint posterior distribution, commonly referred to as Convergence Diagnostic and Output Analysis (CODA), thereby preserving the underlying joint distribution. In addition, the effectiveness of lapatinib in combination with capecitabine is based upon the Kaplan-Meier curve from EMILIA until 72 weeks and hence no uncertainty is assumed around the relative efficacy between this treatment and T-DM1 until beyond 72 weeks.

The uncertainty around the coefficients used to estimate the utility values are based upon the standard errors reported within Lloyd *et al*.[23] No uncertainty is incorporated around the treatment response and AE rates from the trials which are used to generate the utility values, although the latter are implemented within the statistical model by Lloyd *et al.* incorrectly by the manufacturer anyway (see Section 5.2.7). The uncertainty around cost estimates within the model appears to be arbitrary.

Tabled results of the probabilistic sensitivity analysis are not presented within the MS.

The univariate sensitivity analyses presented by the manufacturer compared T-DM1 with capecitabine only. Given that lapatinib in combination with capecitabine is neither dominated nor extendedly dominated by any of the comparators, it would be more appropriate to present a comparison of T-DM1 with this intervention. Moreover, all comparisons should have been included within the one way sensitivity analysis because the appropriate incremental comparison may change for each analysis. Thus, the impact upon the model results of changing parameters within the model is not well described by the manufacturer's analysis. The results of the manufacturer's univariate sensitivity analyses are shown within Table 19. Key results of the sensitivity analyses have been recalculated by the ERG (see Sections 5.3 and 6).

Table 19: Results of the univariate sensitivity analysis presented within the MS

	Base case value (BCV)	High value	Low value
Costs			
PFS supportive care cost (weekly)	£43.45	£86.91 (BCV x2)	£21.73 (BCV x 0.5)
ICER		£113,003	£110,142
Progressed disease supportive care cost (weekly)	£63.08	£123.99 (BCV x2)	£31.54 (BCV x 0.5)
ICER		£113,158	£110,064
Post-progression lines (2L/3L) of treatment duration	4.3 months (18.63 weeks)	8.6 months (37.27 weeks)	2.15 months (9.32 weeks)
ICER		£111,737	£110,775
Treatment dose	Planned treatment dose	Actual treatment dose observed in the trial for all whole duration of progression-free survival	
ICER		£111,871	
Drug costs	Including wastage	Excluding wastage (full vial sharing) for all drugs (except vinorelbine as generic and assumed to be made up through compounders)	
ICER		£108,082	
Outcomes (Results from the MTC)			
PFS HR: Tra+Cap vs T-DM1	0.68	0.50 (95% CI)	0.91 (95% CI)
ICER		£111,069	£111,132
PFS HR: Cap vs T-DM1	0.39	0.29 (95% CI)	0.55 (95% CI)
ICER		£108,700	£115,191
OS HR: Tra+Cap vs T-DM1	0.68	0.46 (95% CI)	0.98 (95% CI)
ICER		£111,205	£110,998
OS HR: Cap vs T-DM1	0.55	0.41 (95% CI)	0.75 (95% CI)
ICER		£89,965	£165,517
PFS HR: Tra+Cap vs T-DM1	0.68	0.54 (excluding CEREBEL and Martin *et al.*)	
ICER		£111,075	
PFS HR: Cap vs T-DM1	0.39	0.35 (excluding CEREBEL and Martin *et al.*)	
ICER		£110,123	
OS HR: Tra+Cap vs T-DM1	0.68	0.58 (excluding CEREBEL and Martin *et al.*)	
ICER		£111,140	
OS HR: Cap vs T-DM1	0.55	0.52 (excluding CEREBEL and Martin *et al.*)	
ICER		£105,788	

	Base case value (BCV)	High value	Low value
Progression-free utility	Different values in arms	0.74 (same values as Lap and Cap arm in all arms)	
ICER		£118,617	
Progression-free utility T-DM1	0.78	0.71 (T-DM1 PFS utilty from TH3RESA trial)	
		£123,257	
Progressed utility	0.50	0.70 (BCV +0.2)	0.30 (BCV -0.2)
ICER		£98,511	£126,660
Progression free utility T-DM1	0.78	0.98 (BCV +0.2)	0.58 (BCV -0.2)
ICER		£94,909	£179,337
Progression free utility Lap + cap	0.74	0.94 (BCV +0.2)	0.54 (BCV -0.2)
ICER		£111,095	£111,095
Progression free utility Tra + cap/Tra + vin	0.73	0.93 (BCV +0.2)	0.53 (BCV -0.2)
ICER		£111,095	£111,095
Progression free utility Cap/Vin	0.72	0.92 (BCV +0.2)	0.52 (BCV -0.2)
ICER		£123,971	£100,371
Parametric functions			
PFS	Kaplan-Meier data with log-normal tail	Kaplan-Meier data with other parametric tails (1) Weibull (2) Exponential (3) Log-logistic (4) Gamma (5) Piecewise exponential tail (one piece)	
ICER		(1) £100,365 (2) £106,672 (3) £114,826 (4) £110,015 (5) £106,211	
OS	Gamma distribution	Other parametric distributions (1) Weibull (2) Log-logistic (3) Log-normal (4) KM data with piecewise exponential tail (one piece) (5) KM data with piecewise exponential tail (two pieces)	
ICER		(1) £151,208 (2) £115,020 (3) £111,004 (4) £138,286 (5) £153,319	
Other			
Cost discount rate	3.5%	6%	0%

	Base case value (BCV)	High value	Low value
ICER		£108,305	£115,586
Health outcomes discount rate	3.5%	6%	0%
ICER		£118,396	£100,816
Health and cost discount rates	3.5% both arms	6% both arms	0% both arms
ICER		£115,413	£104,873
Time horizon	10	15	5
ICER		£107,657	£133,103

5.2.11 Model validation and face validity check

The MS reports that the MTC and economic evaluation have been validated through consultation with clinical and modelling experts and through comparison with external data. In particular, the extrapolated trial data was compared with 10 year registry data of people with HER2-positive MBC. Although these patients had not received T-DM1, it provided information about the expected shape of the survival curves. The manufacturer also reports undertaking internal validity checks.

The ERG externally validated the model through consultation with clinical experts and comparison with existing data sources. Internal validation involved checking the model formula and code and checking face validity of model results both within the base case and the sensitivity analysis.

5.3 Exploratory and sensitivity analyses undertaken by the ERG

5.3.1 The ERG's suggested base case

Based upon the critique of the manufacturer's economic model, the ERG have identified two key errors in implementation and four key assumptions which are methodologically weak which have been revised for the ERG's base case. Thus the ERG's suggested base case includes:

1) *Correcting the cost of AEs.*

As described within Section 5.2.8.5, the model was intended to include the cost of treatment associated with those grade 3 and 4 adverse events with over 2% incidence in either treatment arm of EMILIA; however there are some coding errors which mean that only costs associated with diarrhoea and fatigue are included. This is corrected within the ERG's base case. In addition, within the ERG's base case the total weekly costs associated with AEs have been multiplied by the proportion of patients in PFS rather than the proportion of patients on treatment since the weekly cost of AEs has already taken into account time on treatment.

2) *Correcting the code for the cost of administration of trastuzumab in combination with vinorelbine.*

As described within Section 5.2.8.2, the code for the cost of administration of trastuzumab in combination with vinorelbine resulted in some negative weekly costs. Trastuzumab is administered every three weeks whilst vinorelbine is administered on a weekly basis. The submitted model code multiplied the cost of administration of trastuzumab in combination with vinorelbine by the proportion of patients in PFS during that cycle and subtracted half of the pharmacy costs every three weeks. The ERG's base case multiplies the cost of administration of trastuzumab in combination with vinorelbine by the proportion of patients in PFS every three weeks and multiplies the cost of administration of vinorelbine monotherapy by the proportion of patients in PFS during the remaining two out of every three weeks.

3) *Using hazard ratios from the ERG's random effects MTC.*

As described within Section 4.5, the ERG has undertaken additional analysis to produce results from a random effects MTC to account for between-study variability. The HRs from this analysis have been employed within the ERG base case.

4) *Calculating weekly costs in the progressed state independently for each treatment option.*

As described within Section 5.2.8.4, within the manufacturer's model the total weighted cost of treatment in the progressed state is divided by the average time in the progressed disease health state across all interventions. Within the ERG's base case, the total weighted cost associated with post-progression has been divided by the mean time in that state *for that treatment option*, so that the total cost associated with post-progression is the same for each treatment option to be consistent with the assumptions described by the manufacturer.

5) *A 15 year time horizon (rather than 10 years).*

As described within Section 5.2.5, patients are followed over 10 years within the MS base case; however this was increased to 15 years within the clarification process since up to 3% of patients remain alive at 10 years within the model. A 15 year time horizon more fully captures the differences between costs and outcomes of the interventions.

6) *Incorporation of variation in patient's weight and body surface area to calculate the dosage of T-DM1, capecitabine and trastuzumab.*

As described within Section 5.2.8.1, the mean body weight and surface area of patients within EMILIA is used by the manufacturer to estimate drug costs. Ignoring variability is likely to lead to inaccurate estimates. An accurate estimate of the dosage of T-DM1 could be derived from the patient-level data from EMILIA which would provide distributions of weight for the patient population of

interest. Within their clarification letter, the manufacturer's suggested that undertaking this additional analysis around the cost of the drug "given the magnitude of the base-case ICERs… …appears to be a second order issue". In the absence of patient-level data and given that no data have been identified by the ERG around the weight distribution of adults with HER2-positive metastatic breast cancer, the dose of T-DM1 has been recalculated using the mean weight from the EMILIA trial and assuming a normal distribution with a standard deviation of 15.3. This is based upon a study reporting the body surface area distribution of adult cancer patients in the UK,[30] with the weight distribution provided via personal correspondence. The mean weight in this study is similar to those patients in the EMILIA trial (71.8kg versus 70.1kg). It is also consistent with the weight distribution of the general population of adults in England from the Health Survey for England dataset. The cost of trastuzumab has been recalculated using the same method for the ERG's base case.

The cost of capecitabine has also been recalculated using the distribution of body surface area, based upon the paper by Sacco *et al*.[30]

For both T-DM1 and capecitabine the actual dose according to patient weight from the EMILIA trial was used rather than the planned dose for the ERG's base case. This is important because the effectiveness estimates are based upon the actual dose provided and the actual dose of capecitabine was substantially lower than the planned dose.

Table 20 shows the three-weekly cost of each treatment used within the manufacturer's base case and within the ERG's base case.

Table 20: Three-weekly cost of each treatment

Treatment	**Manufacturer's base case cost**	**ERG's base case cost**
T-DM1	£4267	£4410
Lapatinib	£1206	£1206
Capecitabine (1000mg/kg)	£223	£188
Capecitabine (1250mg/kg)	£285	£278
Trastuzumab Initial dose Maintenance dose	 £1630 £1222	 £1726 £1345
Vinorelbine	£69	£69

Within the economic model, the calculations for applying the actual dose for T-DM1 and capecitabine for the sensitivity analysis appear to contain an error similar to that reported for estimating the cost of

AEs. According to the MS, the actual treatment dose is the dose observed in the trial for the whole duration of progression-free survival, meaning it will include dose interruptions and treatment discontinuation. Within the model, the cost of the three-weekly dose is recalculated according to the reduced mg per kg provided to patients; however, the proportion of patients remaining on treatment, rather than the proportion in PFS, is multiplied by this reduced cost. Thus, the ERG's base case adjusts the cost of the three-weekly dose using the values in Table 20 above, but uses the proportion in PFS rather than the proportion of patients remaining on treatment.

5.3.2 Probabilistic sensitivity analysis

Section 5.2.10 described substantial limitations with the PSA within the economic model. The ERG believes that the following corrections would need to be undertaken as a minimum to provide reasonable expected results:

- Reanalyse the survival data for deriving the parametric distributions for PFS and OS;
- Use the joint posterior distribution of (log) hazard ratios from the MTC;
- Use informed parameters for the uncertainty around costs and utilities.

Given the substantial resources that would be required in delivering the above, the ERG has focused upon correcting the deterministic base case analysis and undertaking substantial one way sensitivity analysis using the deterministic model to describe the key drivers of the model results rather than producing robust PSA results. It is highly unlikely that such a PSA would reduce the mean ICER for T-DM1 to the £20,000 - £30,000 cost per QALY gained quoted within the NICE Guide to the Methods of Technology Appraisal.[31]

5.3.3 Univariate sensitivity analysis

The ERG have repeated selected univariate sensitivity analyses run by the manufacturer, chosen based upon key areas of uncertainty identified within the ERG's critique of the model. A table describing the rationale for which sensitivity analyses are repeated is provided within Appendix 5. In addition to the sensitivity analysis presented within the MS, based upon the critique of the economic model, the ERG has tested:

1) Assuming no benefit beyond the trial duration by setting PFS and OS of T-DM1 equivalent to that of lapatinib in combination with capecitabine after week 72 and 96 respectively (i.e. when numbers at risk become low in EMILIA);
2) Doubling the weekly cost associated with AEs;
3) The impact of trastuzumab being given in its alternative form as a fixed dose subcutaneous administration by decreasing the cost of trastuzumab (to £1222.20 per three weekly cycle) and its administration (trastuzumab in combination with capecitabine assumed to be same as lapatinib in combination with capecitabine at £147.53 per administration).

5.4 Conclusions of the cost effectiveness section

The MS did not identify any existing economic evaluations of T-DM1. The *de novo* model developed is appropriate for the decision problem defined in the final NICE scope and was generally well described within the report. The model structure is clinically appropriate. Following the clarification process, the manufacturer's reported a deterministic ICER for T-DM1 compared with lapatinib in combination with capecitabine of £167,253, the latter of which is estimated to have an ICER of £39,449 compared with capecitabine alone. The ERG has identified two key errors in implementation and four key assumptions which are methodologically weak which have been revised for the ERG's base case, although these do not impact substantially upon the model results (see Section 6).

The uncertainty around the model parameters for the PSA is inadequately characterised and the PSA results are not tabled within the MS. The one way sensitivity analysis provided by the manufacturer does not establish the robustness of the model results or determine the key drivers of the ICERs because T-DM1 is compared with capecitabine only.

6. IMPACT ON THE ICER OF ADDITIONAL CLINICAL AND ECONOMIC ANALYSES UNDERTAKEN BY THE ERG

6.1 ERG's base case ICER

The ERG's base case ICER is developed in stages in Tables 21 – 26. As in the manufacturer's base case, vinorelbine, trastuzumab in combination with capecitabine, and trastuzumab in combination with vinorelbine are dominated. Whilst the revised drug costs increase the ICER associated with T-DM1 compared with lapatinib in combination with capecitabine, the other changes reduce the ICER, resulting in an incremental cost per QALY of £166,429, which is very similar to that submitted by the manufacturer within the clarification responses. All tables show a full incremental analysis. The results have been presented in order of ascending effectiveness rather than costs, as opposed to the manufacturer's, to avoid changing the order of the interventions within the tables due to the costs of lapatinib in combination with capecitabine and trastuzumab in combination with capecitabine being similar.

Table 21: Manufacturer's base case

Technologies	Totals			Incrementals			
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	ICER (Cost per QALY gained)
Capecitabine	£11,850	1.61	0.89				
Vinorelbine	£16,518	1.61	0.89	£4,668	0.00	0.00	Dominated
Trastuzumab and capecitabine	£35,784	2.24	1.28	£19,266	0.63	0.40	Dominated
Trastuzumab and vinorelbine	£36,662	2.24	1.28	£878	1.61	0.89	Dominated
Lapatinib and capecitabine	£34,227	2.53	1.45	-£2,435	0.29	0.17	£39,449
T-DM1	£111,226	3.16	1.91	£76,999	0.63	0.46	£167,253

Table 22 shows the model results when the resource use for all adverse events with over 2% incidence in either treatment arm of EMILIA is included correctly and the weekly cost of AEs is multiplied by the proportion of patients in PFS rather than the proportion of patients on treatment. This does not impact substantially upon the base case results.

Table 22: Correcting the cost of AEs

Technologies	Totals			Incrementals			
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	ICER (Cost per QALY gained)
Capecitabine	£11,933	1.61	0.89				
Vinorelbine	£16,601	1.61	0.89	£4,668	0.00	0.00	Dominated
Trastuzumab and capecitabine	£35,945	2.24	1.28	£19,344	0.63	0.40	Dominated
Trastuzumab and vinorelbine	£36,823	2.24	1.28	£878	1.61	0.89	Dominated
Lapatinib and capecitabine	£34,388	2.53	1.45	-£2,434	0.29	0.17	£39,588
T-DM1	£111,385	3.16	1.91	£76,997	0.63	0.46	£167,246

Table 23 shows the model results when the cost of AEs is corrected as above, in combination with correcting the code for the cost of administration of trastuzumab in combination with vinorelbine and calculating weekly costs in the progressed disease state independently for each treatment option. Again, this does not impact substantially upon the base case results.

Table 23: Correcting the code for the cost of administration of trastuzumab in combination with vinorelbine and weekly costs in the progressed state

Technologies	Totals			Incrementals			
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	ICER (Cost per QALY gained)
Capecitabine	£12,242	1.61	0.89				
Vinorelbine	£16,910	1.61	0.89	£4,668	0.00	0.00	Dominated
Trastuzumab and capecitabine	£35,970	2.24	1.28	£19,060	0.63	0.40	Dominated
Trastuzumab and vinorelbine	£42,409	2.24	1.28	£6,439	1.61	0.89	Dominated
Lapatinib and capecitabine	£34,180	2.53	1.45	-£8,229	0.29	0.17	£38,676
T-DM1	£110,926	3.16	1.91	£76,745	0.63	0.46	£166,701

Table 24 shows the model results following the above changes in combination with using the HRs reported from the ERG's random effects MTC. This increases the ICER for lapatinib in combination with capecitabine compared with capecitabine monotherapy from £38,676 to £52,884.

Table 24: Applying the HRs from the ERG's random effects MTC

Technologies	Totals			Incrementals			
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	ICER (Cost per QALY gained)
Capecitabine	£13,672	1.93	1.07				
Vinorelbine	£19,520	1.93	1.07	£5,848	0.00	0.00	Dominated
Trastuzumab and capecitabine	£37,531	2.27	1.31	£18,010	0.34	0.24	Dominated
Trastuzumab and vinorelbine	£44,375	2.27	1.31	£6,844	1.93	1.07	Dominated
Lapatinib and capecitabine	£34,180	2.53	1.45	-£10,195	0.26	0.14	£52,884
T-DM1	£110,926	3.16	1.91	£76,745	0.63	0.46	£166,701

Table 25 shows the model results following the above changes in combination with employing a 15 year time horizon rather than a 10 year time horizon. This decreases the ICER for T-DM1 compared with lapatinib in combination with capecitabine by around £7,000.

Table 25: A 15 year time horizon

Technologies	Totals			Incrementals			
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	ICER (Cost per QALY gained)
Capecitabine	£13,701	1.94	1.07				
Vinorelbine	£19,553	1.94	1.07	£5,852	0.00	0.00	Dominated
Trastuzumab and capecitabine	£37,662	2.28	1.32	£18,109	0.35	0.25	Dominated
Trastuzumab and vinorelbine	£44,526	2.28	1.32	£6,864	1.94	1.07	Dominated
Lapatinib and capecitabine	£34,349	2.56	1.47	-£10,178	0.27	0.15	£51,760
T-DM1	£111,942	3.24	1.95	£77,593	0.68	0.49	£159,486

Table 26 shows the model results following all of the above changes in combination with incorporating the variation in patients' weight and body surface area to calculate the dosage of T-DM1, capecitabine and trastuzumab. This increases the ICER for T-DM1 compared with lapatinib in combination with capecitabine by around £7,000.

Table 26: Incorporation of variation in patients' weight and body surface area to calculate drug dosage

Technologies	Totals			Incrementals			
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	ICER (Cost per QALY gained)
Capecitabine	£13,628	1.94	1.07				
Vinorelbine	£19,540	1.94	1.07	£5,912	0.00	0.00	Dominated
Trastuzumab and capecitabine	£39,249	2.28	1.32	£19,709	0.35	0.25	Dominated
Trastuzumab and vinorelbine	£46,211	2.28	1.32	£6,962	1.94	1.07	Dominated
Lapatinib and capecitabine	£33,821	2.56	1.47	-£12,390	0.27	0.15	£50,620
T-DM1	£114,792	3.24	1.95	£80,971	0.68	0.49	£166,429

6.2 Univariate sensitivity analysis

The incremental cost per QALYs from the ERG's univariate sensitivity analyses are shown within Table 27 below. Vinorelbine monotherapy and trastuzumab in combination with vinorelbine have been removed from the table for ease of reading, since these treatment options are always dominated by capecitabine monotherapy and trastuzumab in combination with capecitabine respectively.

Table 27: ERG's one way sensitivity analysis

Analysis (BCV= Base case value)	**Capecitabine**	**Trastuzumab and capecitabine**	**Lapatinib and capecitabine**	**T-DM1**
Base case	-	Dominated	£50,620	£166,429
PFS supportive care cost £86.91 (BCVx2) £21.73 (BCVx0.5)	-	 Dominated Dominated	 £54,146 £48,857	 £169,603 £164,842
Post-progression supportive care cost BCVx2: variable between comparators from £117-£129	-	 Dominated	 £52,228	 £167,731
Treatment dose Including wastage - planned Excluding wastage - actual Excluding wastage - planned	-	 Dominated Dominated Dominated	 £51,961 £50,192 £51,760	 £170,762 £149,707 £153,980
Trastatuzumab and capecitabine vs T-DM1 HR PFS 1.48 (Upper CrI)	-	Extendedly dominated	£50,620	£166,429
Trastatuzumab and capecitabine vs T-DM1 HR OS 1.60 (Upper CrI)	-	£20,786	Extendedly dominated	Dominated
Capecitabine vs T-DM1 HR PFS 0.89 (Upper CrI)	-	Dominated	£57,962	£166,429
Capecitabine vs T-DM1 HR OS 1.30 (Upper CrI)	Dominates comparators	Dominated	Dominated	Dominated
PFS utility: Same values as lapatinib and capecitabine in all arms TH3RESA trial (0.71 T-DM1, 0.69 comparators)	 - -	 Dominated Dominated	 £51,727 £54,102	 £185,623 £183,966
Progressed utility 0.7 (BCV +0.2)	-	Dominated	£44,226	£148,983
PFS extrapolation KM+Weibull tail Weibull	 - -	 Dominated Dominated	 £46,646 £47,110	 £147,528 £148,690
OS extrapolation KM+Weibull tail Weibull	 - -	 Dominated Dominated	 £91,952 £90,025	 £191,776 £199,154
PFS & OS extrapolation - Weibull	-	Dominated	£89,433	£181,263

Analysis (BCV= Base case value)	Capecitabine	Trastuzumab and capecitabine	Lapatinib and capecitabine	T-DM1
Discount rate (costs & outcomes) 6% 0%	 - -	 Dominated Dominated	 £52,852 £47,412	 £174,951 £154,012
Time horizon - 5 years	-	Dominated	£60,284	£217,513
PFS & OS of T-DM1 equivalent to lapatinib and capecitabine after week 72 and 96 respectively	-	Dominated	£50,620	£449,554
Cost of AEs (BCVx2)	-	Dominated	£51,146	£165,858
Fixed dose subcutaneous trastuzumab administration	-	Dominated	£50,620	£166,429

All of these analyses result in an incremental cost per QALY gained for T-DM1 compared with lapatinib in combination with capecitabine in excess of £147,000. The ICER associated with lapatinib in combination with capecitabine compared with capecitabine does not fall below £44,000.

The ICER for T-DM1 compared with lapatinib in combination with capecitabine is reduced by more than 10% by:

- excluding wastage from the drug costs;
- increasing the utility associated with progressed disease from 0.5 to 0.7;
- fitting a weibull distribution to the tail of the PFS curve rather than a lognormal distribution.

For T-DM1 compared with lapatinib in combination with capecitabine, the ICER is increased by more than 10% by:

- assuming consistent utilities across treatment options in PFS/ using utility values from interim results of TH3RESA;
- fitting a weibull distribution for OS rather than a gamma distribution;
- reducing the time horizon to 5 years;
- setting PFS and OS for T-DM1 equivalent to lapatinib in combination with capecitabine after weeks 72 and 96 respectively.

The MTC results show that the comparative effectiveness between treatment options is highly uncertain. If any of the comparators were to have equivalent overall survival impacts to T-DM1 then they would dominate T-DM1 due to the higher acquisition costs associated with T-DM1.

Assuming a fixed dose subcutaneous trastuzumab administration does not reduce the cost of trastuzumab in combination with capecitabine below that of lapatinib in combination with capecitabine, meaning that trastuzumab in combination with capecitabine remains dominated.

7. END OF LIFE

The MS does not propose a case for meeting end of life criteria.

To meet NICE end of life criteria all of the below must be satisfied:

1) The treatment is indicated for patients with a short life expectancy, normally less than 24 months and;
2) There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared to current NHS treatment, and;
3) The treatment is licensed or otherwise indicated, for small patient populations.

Based upon the EMILIA trial, 51.8% of patients remained alive at 24 months and patients had a median overall survival of 25.1 months from treatment initiation with lapatinib in combination with capecitabine. This extended by 5.8 months to a median overall survival of 30.9 months with T-DM1, with 64.7% of patients remaining alive at 24 months. The manufacturer estimated that the eligible population would be approximately 1,290 patients per year. Thus the ERG believes that criteria 2 and 3 for end of life would be met by T-DM1, whilst criterion 1 would not be met.

8. OVERALL CONCLUSIONS

The clinical effectiveness data relevant to the decision problem were taken from two large RCTs, both of which were open-label, but otherwise at low risk of bias, with adverse event data from additional trials. Data from these two RCTs reported a statistically significant advantage in PFS for T-DM1 over lapatinib in combination with capecitabine, and over the treatment of physician's choice. Data also reported a statistically significant advantage in OS and time to symptom worsening for T-DM1 over lapatinib in combination with capecitabine. For T-DM1, the most common grade 3 or greater AEs were thrombocytopenia and hepatotoxicity.

There was a lack of head-to-head comparison with T-DM1 for most comparators in the decision problem. Most of the data were from third-line or later therapy, whereas the MS suggests T-DM1 as second-line treatment, and there were only a few patients with ECOGPS2 from one trial providing data. Within the MTC analysis, allowing for heterogeneity between studies increases the uncertainty about the true treatment effect on OS and PFS. T-DM1 appears to be the best treatment in terms of both OS and PFS. From the ERG's random effects model, T-DM1 is associated with a reduction in the hazard of death of 32% (HR=0.68, 95% CrI [0.37, 1.25]) and a reduction in the hazard of progression or death of 35% (HR=0.65, 95% CrI [0.35, 1.20]) compared to lapatinib in combination with capecitabine.

The *de novo* model developed by the manufacturer is appropriate for the decision problem defined in the final scope and was generally well described within the report. The model structure was considered to be clinically appropriate. Following the clarification process, the manufacturer's reported a deterministic ICER for T-DM1 compared with lapatinib in combination with capecitabine of £167,253, the latter of which is estimated to have an ICER of £39,449 compared with capecitabine monotherapy. The ERG produced very similar revised base case values of £166,429 and £50,620 respectively. The uncertainty around the model inputs for the PSA was inappropriately characterised within the MS. In addition, the sensitivity analysis provided by the manufacturer does not establish the robustness of the model results or determine the key drivers of the results because T-DM1 is compared with capecitabine only. The deterministic sensitivity analysis undertaken by the ERG suggests that the key drivers of the model results are: the relative OS associated with the interventions; the distribution employed for extrapolation of PFS and OS; whether the treatment effect is assumed to continue beyond the trial data; the utility values associated with PFS and progressed disease; and whether wastage is included within the drug costs. However, the ICER for T-DM1 versus lapatinib in combination with capecitabine did not decrease below £147,000 within any of the one way sensitivity analyses.

8.1 Implications for research

T-DM1 is given to patients until they progress or experience an adverse event within current trials. Further trial research comparing different treatment duration options could be worthwhile, particularly given the high current acquisition costs of the drug.

Within randomised controlled trials, T-DM1 is compared only with lapatinib in combination with capecitabine. Trials directly comparing T-DM1 with capecitabine monotherapy or trastuzumab in combination with capecitabine may be informative. However, given current drug acquisition costs and currently acceptable cost-effectiveness acceptability thresholds,[31] it is likely that further data collection comparing T-DM1 with alternative treatments within this indication would not represent value for money, although a formal value of information analysis has not been undertaken.

9. REFERENCES

1. NICE. NICE clinical guideline 81: Advanced Breast Cancer, diagnosis and treatment. 2009; available from http://www.nice.org.uk/CG81 (accessed Jan. 2014).

2. European Medicines Agency. 2014; available from http://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CCoQFjAA&url=http%3A%2F%2Fwww.ema.europa.eu%2Fdocs%2Fen_GB%2Fdocument_library%2FEPAR_-_Product_Information%2Fhuman%2F002389%2FWC500158593.pdf&ei=5iwHU_ThC8mX7QbxhYDoDQ&usg=AFQjCNEBkyJodqlp96kv0Dgj9gVxnDy-vA&bvm=bv.61725948,d.ZGU&cad=rja (accessed Jan 2014)

3. Hurvitz, S.A., Dirix, L., Kocsis, J., Bianchi, G.V., Lu, J., Vinholes, J. et al. Phase II randomized study of trastuzumab emtansine versus trastuzumab plus docetaxel in patients with human epidermal growth factor receptor 2-positive metastatic breast cancer.[Erratum appears in J Clin Oncol. 2013 Aug 10;31(23):2977]. *Journal of Clinical Oncology* 2013; 31(9):1157-1163.

4. Verma, S., Miles, D., Gianni, L., Krop, I.E., Welslau, M., Baselga, J. et al. Trastuzumab emtansine for HER2-positive advanced breast cancer.[Erratum appears in N Engl J Med. 2013 Jun 20;368(25):2442]. *New England Journal of Medicine* 2012; 367(19):1783-1791.

5. Wildiers. TDM1 for HER2-Positive MBC: Primary Results From TH3RESA, a Phase 3 Study of T-DM1 vs Treatment of Physician's Choice. *European Society for Medical Oncology 2013 oral presentation* 2013.

6. Welslau M et al. Patient-Reported Outcomes From EMILIA, a Randomized Phase 3 Study of Trastuzumab Emtansine (T-DM1) Versus Capecitabine and Lapatinib in Human Epidermal Growth Factor Receptor 2-Positive Locally Advanced or Metastatic Breast Cancer. *Cancer 2013; Published ahead of print DOI: 10.1002/cncr.28465* 2013.

7. Centre for Reviews and Dissemination. Systematic reviews: CRD's guidance for undertaking reviews in health care. 2009; 3rd edition. York, University of York.

8. US National Institutes of Health. Clinical Trials Registry. 2014; available from http://clinicaltrials.gov/

9. Masuda, N., Ito, Y., Takao, S., Doihara, H., Rai, Y., Horiguchi, J. et al. A Multicenter Phase 2 Study (JO22997) Evaluating the Efficacy and Safety of Trastuzumab Emtansine in Japanese Patients With Heavily Pretreated HER2-Positive Metastatic Breast Cancer. Poster Session 5 - Treatment: HER2-Targeted Therapy. *Cancer Research* 2012; 72(24):Supplement 3.

10. Dieras V et al. Trastuzumab Emtansine (T-DM1) in Human Epidermal Growth Factor Receptor 2-Positive Metastatic Breast Cancer: Pooled Safety Analysis. *San Antonio Breast Cancer Symposium 2012; Abstract P5-18-06* 2014.

11. Burris, H.A., III, Hurwitz, H.I., Dees, E.C., Dowlati, A., Blackwell, K.L., O'Neil, B. et al. Phase I safety, pharmacokinetics, and clinical activity study of lapatinib (GW572016), a reversible dual inhibitor of epidermal growth factor receptor tyrosine kinases, in heavily pretreated patients with metastatic carcinomas. *J Clin Oncol* 2005; 23(23):5305-5313.

12. Blackwell, K.L., Burstein, H.J., Storniolo, A.M., Rugo, H.S., Sledge, G., Aktan, G. et al. Overall survival benefit with lapatinib in combination with trastuzumab for patients with human epidermal growth factor receptor 2-positive metastatic breast cancer: final results from the EGF104900 Study. *J Clin Oncol* 2012; 30(21):2585-2592.

13. Jerusalem, G., Fasolo, A., Dieras, V., Cardoso, F., Bergh, J., Vittori, L. et al. Phase I trial of oral mTOR inhibitor everolimus in combination with trastuzumab and vinorelbine in pre-treated patients with HER2-overexpressing metastatic breast cancer. *Breast Cancer Res Treat* 2011; 125(2):447-455.

14. Andre, F., Campone, M., O'Regan, R., Manlius, C., Massacesi, C., Sahmoud, T. et al. Phase I study of everolimus plus weekly paclitaxel and trastuzumab in patients with metastatic breast cancer pretreated with trastuzumab. *J Clin Oncol* 2010; 28(34):5110-5115.

15. von, M.G., Schwedler, K., Schmidt, M., Barinoff, J., Mundhenke, C., Cufer, T. et al. Trastuzumab beyond progression: overall survival analysis of the GBG 26/BIG 3-05 phase III study in HER2-positive breast cancer. *Eur J Cancer* 2011; 47(15):2273-2281.

16. Cameron, D., Casey, M., Press, M., Lindquist, D., Pienkowski, T., Romieu, C.G. et al. A phase III randomized comparison of lapatinib plus capecitabine versus capecitabine alone in women with advanced breast cancer that has progressed on trastuzumab: updated efficacy and biomarker analyses. *Breast Cancer Res Treat* 2008; 112(3):533-543.

17. S5-7: A Phase 2, Randomized, Open-Label, Study of Neratinib (HKI-272) vs Lapatinib Plus Capecitabine for 2nd/3rd-Line Treatment of HER2+ Locally Advanced or Metastatic Breast Cancer. Cancer Research; 2011.

18. CEREBEL (EGF111438): An open label randomized Phase III study comparing the incidence of CNS metastases in patients with HER2+ metastatic breast cancer, treated with lapatinib plus capecitabine versus tratuzumab plus capecitabine. European Society of Medical Oncology; 2012.

19. Bucher, H.C., Guyatt, G.H., Griffith, L.E., Walter, S.D. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* 1997; 50(6):683-691.

20. Dias S., Welton N.J., Sutton A.J., Ades A.E. NICE Technical Suport Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials. 2011; available from http://www.nicedsu.org.uk/Evidence-Synthesis-TSD-series%282391675%29.htm (accessed Jan. 2014).

21. Dias S., Sutton A.J., Welton N.J., Ades A.E. NICE DSU Technical Support Document 3: Heterogeneity: Subgroups, meta-regression, bias and bias-adjustment. 2011; available from http://www.nicedsu.org.uk/Evidence-Synthesis-TSD-series%282391675%29.htm (accessed Jan. 2014).

22. Latimer N. NICE DSU Technical Support Document 14: Survival analysis for ecoomic evaluations alongside clinical trials - extrapolation with patient-level data. *Report by the Decision Support Unit* 2013.

23. Lloyd, A., Nafees, B., Narewska, J., Dewilde, S., Watkins, J. Health state utilities for metastatic breast cancer. *Br J Cancer* 2006; 95(6):683-690.

24. Peasgood, T., Ward, S.E., Brazier, J. Health-state utility values in breast cancer. *Expert Rev Pharmacoecon Outcomes Res* 2010; 10(5):553-566.

25. Ara, R., Brazier, J.E. Populating an economic model with health state utility values: moving toward better practice. *Value Health* 2010; 13(5):509-518.

26. NICE. Breast cancer (HER2 positive, metastatic) - pertuzumab (with trastuzumab and docetaxel) [ID523]. 2014; available from http://guidance.nice.org.uk/index.jsp?action=byId&o=13815 (accessed Jan. 2014).

27. British Medical Association, R.P.S.o.G.B. British National Formulary No. 65. BMA, RPSGB, London; 2011.

28. National Institute for Health and Care Excellence. Evidence summary: new medicine: Early and metastatic HER2-positive breast cancer: subcutaneous trastuzumab. 2013; available from http://publications.nice.org.uk/esnm13-early-and-metastatic-her2-positive-breast-cancer-subcutaneous-trastuzumab-esnm13 (accessed Feb. 2014).

29. Curtis L. PSSRU Unit Costs of Health and Social Care 2012. 2012. University of Kent, Personal Social Services Research Unit.

30. Sacco, J.J., Botten, J., Macbeth, F., Bagust, A., Clark, P. The average body surface area of adult cancer patients in the UK: a multicentre retrospective study. *PLoS One* 2010; 5(1):e8933.

31. National Institute for Health and Clinical Excellence. A guide to the methods of technology appraisal. 2013.

10. APPENDICES

Appendix 1: Trials of T-DM1 in HER2-positive LABC or MBC

All these trials were included in the MS

Trial identifier(s)	Relevance to decision problem	Status	Trial design	Data contributing to MS clinical effectiveness section
EMILIA; NCT00829166; TDM4370g; BO21977	Meets NICE scope	Ongoing for OS but completed primary endpoint	RCT phase III	Effectiveness and Adverse event data
TH3RESA; NCT01419197; TDM4997g ; BO25734	Meets NICE scope	Ongoing but PFS primary endpoint reached	RCT phase III	Effectiveness and Adverse event data
NCT00679341; TDM4450g ; BO21976	Outside scope T-DM1 not within licence for most patients (no prior MBC treatment), and comparator outside scope (trastuzumab and docetaxel)	Completed	Phase II RCT	Adverse event data
NCT00943670; TDM4688g	Not comparative study	Completed	Single arm Phase II	Adverse event data
NCT00932373; TDM3569g	Not comparative study	Completed	Phase I	Adverse event data
NCT00509769; TDM4258g	Not comparative study	Completed	Single arm Phase II	Adverse event data
NCT00679211; TDM4374g	Not comparative study	Completed	Single arm Phase II	Adverse event data

Trial identifier(s)	Relevance to decision problem	Status	Trial design	Data contributing to MS clinical effectiveness section
NCT00781612; TDM4529g ; BO25430	Not comparative study	Ongoing	Single arm extension study. Patients from the control arm from Study TDM4450g or other parent study	Adverse event data
JO22997	Not comparative study	Completed	Single arm Phase II	Adverse event data available from abstract
NCT00875979; TDM4373g; BO22495	Not comparative study, T-DM1 plus pertuzumab	Completed	Single arm Phase Ib/II	No
JO22992	Not comparative study, T-DM1 plus pertuzumab	Completed	Single arm Phase Ib	No
JO22591	Not comparative study	Completed	Phase I, dose study	No
KAMILLA; NCT01702571; MO28231	Not comparative study	ongoing	Single arm study (though says phase 3b)	No
NCT01513083; BO25499	Not comparative study	Ongoing	Single arm parallel population (according to hepatic impairment), safety trial	No

Trial identifier(s)	Relevance to decision problem	Status	Trial design	Data contributing to MS clinical effectiveness section
NCT01702558; MO28230 ; TRAX-HER2	Not comparative study, T-DM1 plus capecitabine in MBC or metastatic gastric cancer	Ongoing	Single arm Phase I	No
T-PAS; NCT01120561; TDM4884g; ML01356	Not comparative study	Ongoing, primary endpoint published	Single arm expanded access (i.e. patients who can't be in clinical trial)	No
MARIANNE; NCT01120184; TDM4788g; BO22589	T-DM1 plus pertuzumab or placebo, compared with trastuzumab plus paclitaxel or docetaxel	Ongoing	RCT phase III	No
NCT00951665; TDM4652g	Not comparative study, dose study, T-DM1 plus paclitaxel plus pertuzumab	Ongoing	Phase Ib dose escalation followed by phase IIa extension (T-DM1 plus paclitaxel with or without pertuzumab)	No
NCT00934856; BP22572	Not comparative study, T-DM1 plus pertuzumab plus docetaxel	Ongoing	Single arm Phase Ib/II	No

Trial identifier(s)	Relevance to decision problem	Status	Trial design	Data contributing to MS clinical effectiveness section
NCT00928330; GDC4627g	Not comparative study, T-DM1 plus GDC-0941	Ongoing	Phase I	No

Appendix 2: Effectiveness trial eligibility criteria (taken from Section 6.3.3 of the MS)

EMILIA

Inclusion Criteria

Disease-Specific Criteria

1. Prospective centrally confirmed HER2-positive (i.e., immunohistochemistry [IHC] 3 + and/or gene-amplified by FISH). Both IHC and FISH assays will be performed; however, only one positive result is required for eligibility. Additional tissue samples will be required to perform all mandatory testing (including qRT-PCR).
2. Histologically or cytologically confirmed invasive breast cancer: incurable, unresectable, locally advanced breast cancer previously treated with multimodality therapy or MBC
3. Prior treatment for breast cancer in the adjuvant, unresectable, locally advanced, or metastatic setting must include both:

   A taxane, alone or in combination with another agent, and trastuzumab, alone or in combination with another agent in the adjuvant, unresectable, locally advanced, or metastatic setting
4. Documented progression of incurable unresectable, locally advanced, or metastatic breast cancer, defined by the investigator:

   Progression must occur during or after most recent treatment for locally advanced/MBC or within 6 months after completing adjuvant therapy.
5. Measurable and/or non-measurable disease. Patients with CNS-only disease are excluded.

General Criteria

6. Age ≥ 18 years
7. Cardiac ejection fraction ≥ 50% by either ECHO or MUGA
8. ECOG performance status of 0 or 1
9. Adequate organ function, evidenced by the following laboratory results within 30 days prior to randomisation:
   - Absolute neutrophil count > 1500 cells/mm3
   - Platelet count > 100,000 cells/mm3
   - Haemoglobin > 9.0 g/dL (Patients were allowed to be transfused red blood cells to this level).
   - Albumin ≥ 2.5 g/dL
   - Total bilirubin ≤ 1.5 upper limit of normal (ULN)

- SGOT (aspartate aminotransferase [AST]), SGPT (SGPT (alanine aminotransferase [ALT]), and alkaline phosphatase (ALP) ≤ 2.5 × ULN with the following exception: patients with bone metastases: ALP ≤ 5 × ULN
- Creatinine clearance > 50 mL/min based on Cockroft-Gault glomerular filtration rate (GFR) estimation: (140 − Age) × (weight in kg) × (0.85 if female)/(72 × serum creatinine)
- International normalized ratio (INR) and activated partial thromboplastin time (aPTT) < 1.5 × ULN (unless on therapeutic coagulation)

10. For women of childbearing potential and men with partners of childbearing potential, agreement to use a highly effective, non-hormonal form of contraception. Acceptable forms of contraception should include two of the following:
    - Placement of non-hormonal intrauterine device (IUD)
    - Condom with spermicidal foam/gel/film/cream/suppository
    - Diaphragm or cervical/vault caps with spermicidal foam/gel/film/cream/suppository

    The above contraception is not a requirement in the case of any of the following:
    - Patient is surgically sterilized (i.e., who have undergone surgical sterilisation with a hysterectomy and/or bilateral oophorectomy)
    - Patient has had no menstrual period for 12 consecutive months, or
    - Patient truly abstains from sexual activity

    Contraception use should continue for the duration of the study treatment and for at least 6 months after the last dose of study treatment.

    Periodic abstinence (e.g., calendar ovulation, symptothermal, post-ovulation methods) and withdrawal are not acceptable methods of contraception. Postmenopausal is defined as ≥ 12 months of amenorrhea.

    Specific country requirements will be followed (e.g., in the United Kingdom, women of childbearing potential and male subjects and their partners of childbearing potential must use two methods of contraception [one of which must be a barrier method] for the duration of the study).

EMILIA

Exclusion Criteria

Cancer-Related Criteria

1. History of treatment with T-DM1
2. Prior treatment with lapatinib or capecitabine

3. Peripheral neuropathy of Grade ≥ 3 per National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) Version 3.0
4. History of other malignancy within the previous 5 years, except for appropriately treated carcinoma in situ of the cervix, non-melanoma skin carcinoma, Stage 1 uterine cancer, synchronous or previously diagnosed HER2-positive breast cancer, or cancers with a similar curative outcome as those mentioned above
5. History of receiving any anti-cancer drug/biologic or investigational treatment within 21 days prior to randomisation except hormone therapy, which can be given up to 7 days prior to randomisation; recovery of treatment-related toxicity consistent with other eligibility criteria
6. History of radiation therapy within 14 days of randomisation. The patient must have recovered from any resulting acute toxicity (to Grade ≤ 1) prior to randomisation
7. Brain metastases that are untreated, symptomatic, or require therapy to control symptoms, as well as a history of radiation, surgery, or other therapy, including corticosteroids, to control symptoms from brain metastases within 2 months (60 days) before randomisation

Cardiopulmonary Function

8. History of symptomatic CHF or serious cardiac arrhythmia requiring treatment
9. History of myocardial infarction or unstable angina within 6 months of randomisation
10. Current dyspnoea at rest due to complications of advanced malignancy or requirement for continuous oxygen therapy

General Criteria

11. Current severe, uncontrolled systemic disease (e.g., clinically significant cardiovascular, pulmonary, or metabolic disease)
12. Pregnancy or lactation
13. Currently known active infection with HIV, hepatitis B virus, or hepatitis C virus
14. Presence of conditions that could affect gastrointestinal absorption: malabsorption syndrome, resection of the small bowel or stomach, and ulcerative colitis
15. History of intolerance (such as Grade 3−4 infusion reaction) to trastuzumab
16. Known hypersensitivity to 5-fluorouracil or known dihydropyrimidine dehydrogenase deficiency
17. Current treatment with sorivudine or its chemically related analogs, such as brivudine
18. Assessed by the investigator to be unable or unwilling to comply with the requirements of the protocol (i.e., unable to swallow pills)

TH3RESA

Inclusion Criteria

1. Signed study-specific Informed Consent Form
2. Age ≥ 18 years
3. Histologically or cytologically documented breast cancer
4. Metastatic or unresectable locally advanced/recurrent breast cancer
5. HER2-positive disease documented as ISH-positive and/or 3+ by IHC on previously collected tumour tissue and prospectively confirmed by Sponsor-designated central laboratory prior to study enrollment

   Tumour material made available for confirmatory central laboratory HER2 testing and exploratory biomarker analyses. Both IHC and ISH assays will be performed; however, only one positive result is required for eligibility. For patients with a history of bilateral breast cancer, HER2-positive status must be demonstrated in primary tumours from both breasts or a biopsy from a single metastatic site
6. Disease progression on the last systemic anti-cancer regimen received as defined by the investigator unless they were intolerant

   Patients who were intolerant to their last systemic anti-cancer regimen may be considered eligible if they satisfy all other inclusion criteria. Intolerance is defined as any treatment-related Grade 4 AE, or any treatment-related Grade 2 or 3 AE that is unacceptable to the patient and persists despite standard countermeasures. The reason for intolerance will be fully documented
7. Prior treatment with trastuzumab, lapatinib, and a taxane in any setting (i.e., neoadjuvant, adjuvant, locally advanced, or recurrent/metastatic) and documented disease progression (by investigator assessment) after at least two regimens of HER2-directed therapy in the metastatic or unresectable locally advanced/recurrent setting. Disease progression must have occurred on both trastuzumab and lapatinib containing regimens except where there was intolerance of lapatinib as defined below
8. A minimum of 6 weeks of prior trastuzumab for the treatment of metastatic or unresectable locally advanced/recurrent disease is required
9. Trastuzumab must have been administered as six consecutive weekly doses or as two consecutive doses on a q3w schedule
10. Patients must have had at least 6 weeks of prior exposure in the metastatic (or unresectable locally advanced/recurrent) setting to lapatinib unless they were intolerant of lapatinib

    Intolerance is defined as any treatment-related Grade 4 AE, or any treatment-related Grade 2 or 3 AE that is unacceptable to the patient and persists despite standard countermeasures. The reason for intolerance will be fully documented

Patients who were found to be intolerant to lapatinib can be considered eligible if they experienced disease progression during a single trastuzumab-based regimen in the metastatic (or unresectable locally advanced/recurrent) setting

11. Adequate organ function, as evidenced by the following laboratory results:

    ANC > 1500 cells/mm^3

    Platelet count > 100,000 cells/mm^3

    Haemoglobin > 9.0 g/dL

    Patients are allowed to receive transfused RBC to achieve this level.

    Total bilirubin ≤ 1.5 × ULN, except in patients with previously documented Gilbert's syndrome, in which case the direct bilirubin should be less than or equal to the ULN

    SGOT (AST) and SGPT (ALT) ≤ 2.5 × ULN

    Alkaline phosphatase ≤ 2.5 × ULN

    Patients with hepatic and/or bone metastases: alkaline phosphatase ≤ 5 × ULN

    Serum creatinine < 1.5 × ULN
12. INR < 1.5 × ULN
13. Eastern Cooperative Oncology Group (ECOG) performance status of 0, 1, or 2 (see Appendix B)
14. LVEF ≥ 50% by either ECHO or MUGA
15. Willingness and ability to comply with scheduled visits, treatment plans, laboratory tests, and other study procedures, including completion of PRO measures
16. Negative results of serum pregnancy test for premenopausal women of reproductive capacity and for women < 12 months after entering menopause
17. For women of childbearing potential and men with partners of childbearing potential, agreement by the patient and/or partner to use a highly effective, non-hormonal form of contraception or two effective forms of non-hormonal contraception:

    Acceptable forms of highly effective contraception include the following:

    True abstinence when this is in line with the preferred and usual lifestyle of the patient. Periodic abstinence (e.g., calendar, ovulation, symptothermal post-ovulation methods) and withdrawal are not acceptable methods of contraception

    Male sterilization (with appropriate post-vasectomy documentation of the absence of sperm in the ejaculate). For female patients, the vasectomized male partner should be the sole partner. Tubal ligation is not considered a highly effective contraception

    Acceptable forms of effective contraception include the following:

    Placement of non-hormonal intrauterine device or intrauterine system

    Condom with spermicidal foam/gel/film/cream/suppository

    Occlusive cap (diaphragm or cervical/vault caps) with spermicidal foam/film/cream/suppository

Contraception as described above is not a requirement in the case of any of the following:

The male patient or male partner of a female patient is surgically sterilized

The female patient is ≥ 45 years of age and is postmenopausal (has had no menstrual period for at least 12 consecutive months)

The female patient has undergone hysterectomy and/or bilateral oophorectomy

Contraception use should continue for the duration of the study treatment and for at least 6 months after the last dose of study treatment

TH3RESA

Exclusion Criteria

Cancer-Related Criteria

1. Chemotherapy ≤ 21 days before first study treatment
2. Trastuzumab ≤ 21 days before first study treatment
3. Lapatinib ≤ 14 days before first study treatment
4. Hormone therapy ≤ 7 days before first study treatment
5. Investigational therapy or any other therapy ≤ 28 days before first study treatment
6. Prior enrollment in a T-DM1−containing study, regardless of whether the patient received prior T-DM1
7. Previous radiotherapy for the treatment of unresectable, locally advanced/recurrent or metastatic breast cancer is not allowed if:

   The last fraction of radiotherapy has been administered within 14 days prior to randomisation

   The patient has not recovered from any resulting acute toxicity (to Grade ≤ 1) prior to randomisation

   Brain metastases that are untreated or symptomatic, or require any radiation, surgery, or corticosteroid therapy to control symptoms from brain metastases within 1 month of randomisation

   For patients with newly diagnosed brain metastases or unequivocal progression of brain metastases on screening scans, prior localized treatment (i.e., surgery, radiosurgery, and/or whole brain radiotherapy) is required
8. History of intolerance (including Grade 3 or 4 infusion reaction) or hypersensitivity to trastuzumab or murine proteins
9. History of exposure to the following cumulative doses of anthracyclines: Doxorubicin or liposomal doxorubicin > 500 mg/m^{2}Epirubicin> 900 mg/m^2, Mitoxantrone> 120 mg/m^2

   If another anthracycline, or more than one anthracycline, has been used, the cumulative dose must not exceed the equivalent of 500 mg/m^2 doxorubicin
10. Current peripheral neuropathy of Grade ≥ 3 per the NCI CTCAE v4.0

11. History of other malignancy within the last 5 years, except for appropriately treated carcinoma in situ of the cervix, non-melanoma skin carcinoma, Stage I uterine cancer, or other cancers with a similar outcome as those mentioned above

Cardiopulmonary Function Criteria

12. Current unstable ventricular arrhythmia requiring treatment
13. History of symptomatic CHF (New York Heart Association [NYHA] Classes II−IV)
14. History of myocardial infarction or unstable angina within 6 months of enrolment
15. History of a decrease in LVEF to < 40% or symptomatic CHF with previous trastuzumab treatment
16. Severe dyspnoea at rest due to complications of advanced malignancy or requiring current continuous oxygen therapy

General Criteria

17. Current severe, uncontrolled systemic disease (e.g., clinically significant cardiovascular, pulmonary, or metabolic disease)
18. Major surgical procedure or significant traumatic injury within 28 days before enrollment or anticipation of the need for major surgery during the course of study treatment
19. Current pregnancy or lactation
20. Current known active infection with HIV, hepatitis B, and/or hepatitis C virus

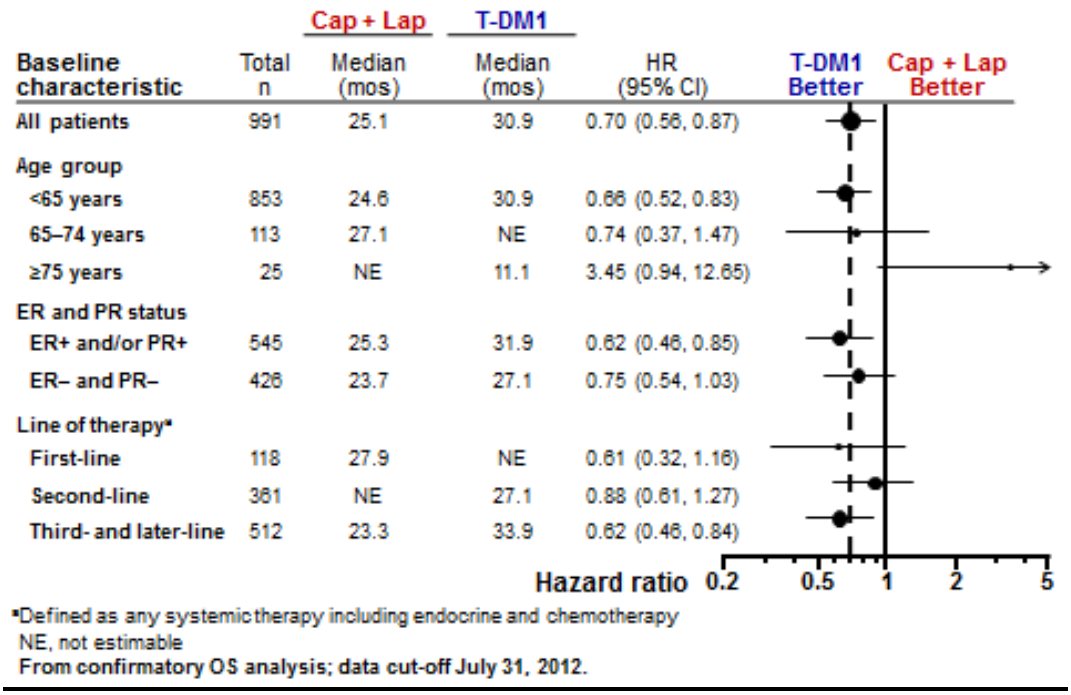
    For patients who are known carriers of hepatitis B virus (HBV), active hepatitis B infection must be ruled out based on negative serologic testing and/or determination of HBV DNA viral load per local guidelines

Assessed by the investigator to be unable or unwilling to comply with the requirements of the protocol

Appendix 3: Subgroup analyses from EMILIA and TH3RESA

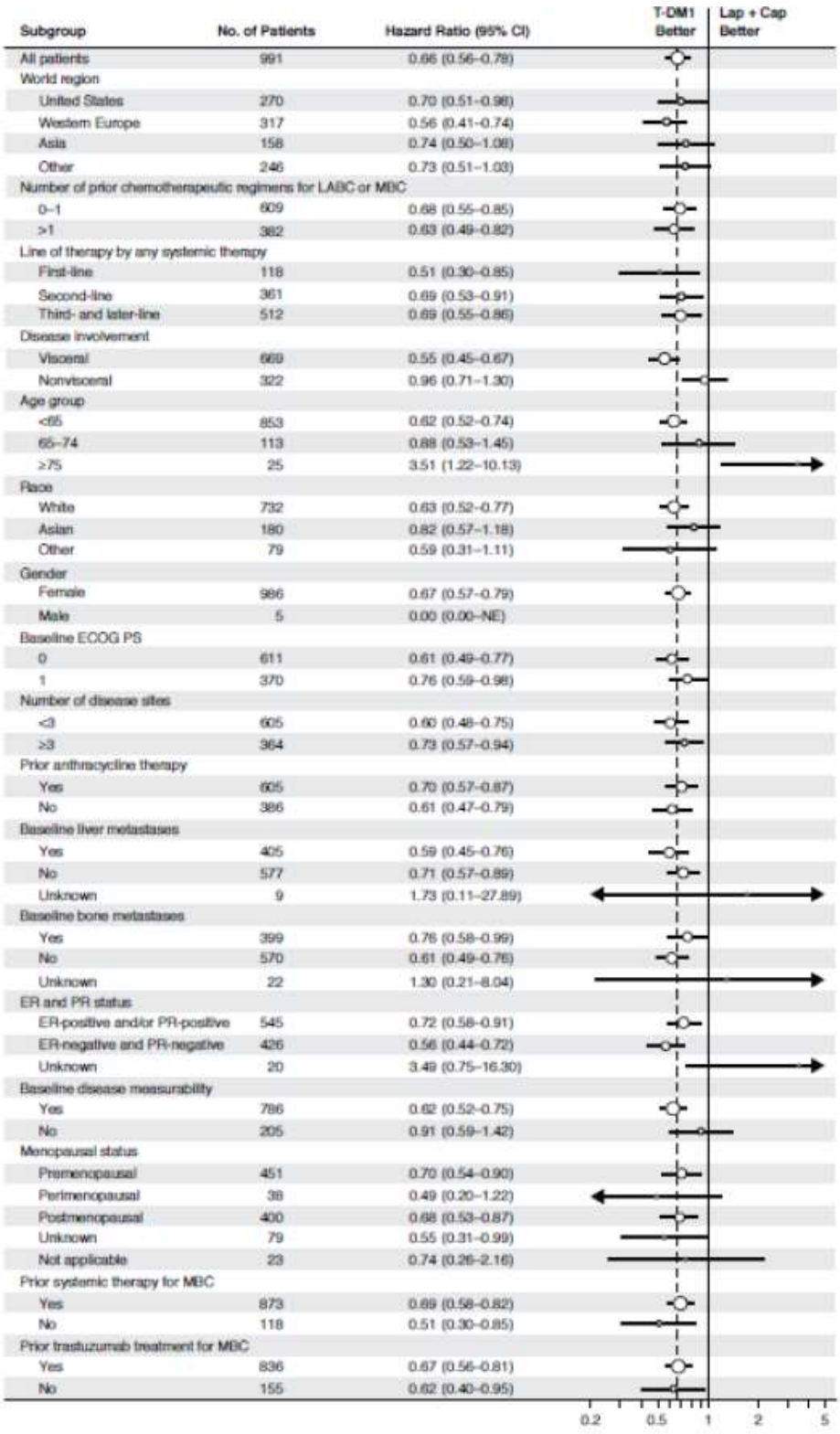
OS results: EMILIA (Figure taken directly from MS)

Overall Survival Subgroup Analyses



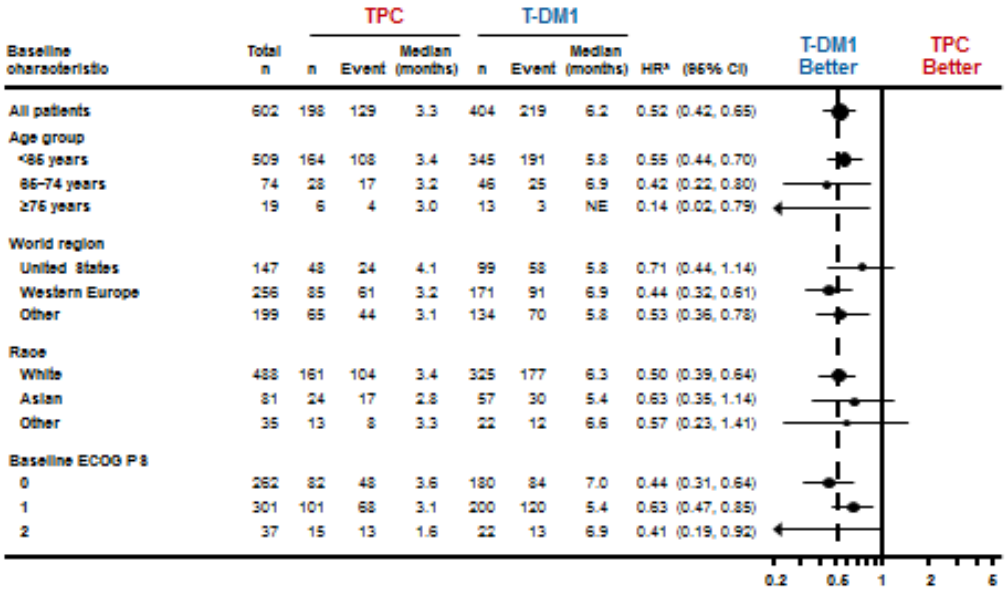
*Defined as any systemic therapy including endocrine and chemotherapy
NE, not estimable
From confirmatory OS analysis; data cut-off July 31, 2012.

PFS results: EMILIA (Figure taken directly from MS)



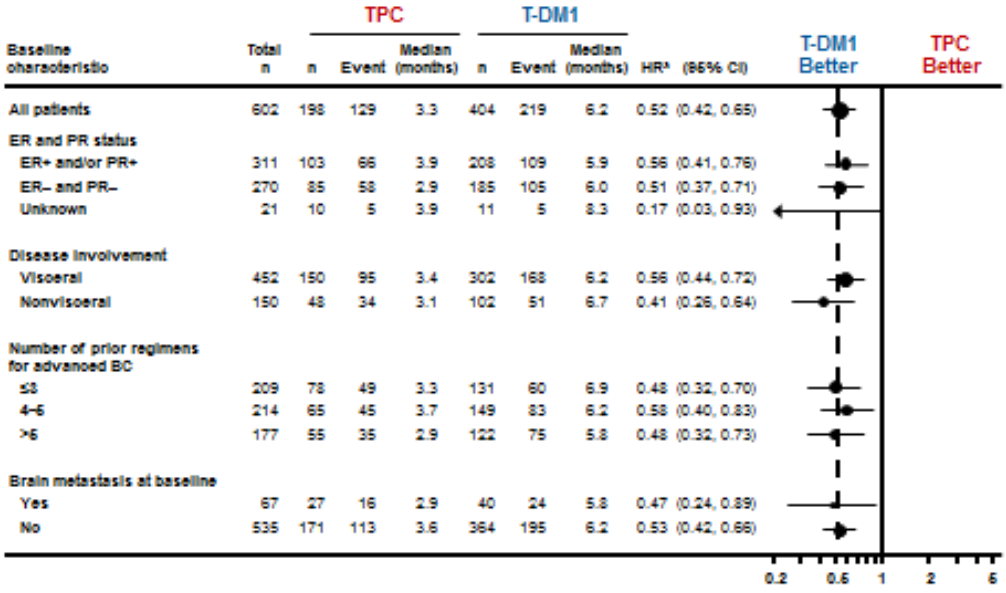
PFS results: TH3RESA (Figures taken directly from MS)

PFS Subgroup Analyses (1)
By Investigator Assessment



* Unstratified HR
NE, not estimable

PFS Subgroup Analyses (2)
By Investigator Assessment



* Unstratified HR

Appendix 4: Table of Common AEs in the pooled analysis of 884 patients treated with single-agent T-DM1 (taken directly from the MS clarification response)

Table 9 Common Adverse Events in Studies of Trastuzumab Emtansine Alone or in Combination with Pertuzumab in Patients with HER2-Positive MBC

Common Adverse Events[a]	T-DM1 N=884	T-DM1 + Pertuzumab N=87
BODY SYSTEM		
Preferred Term		
GENERAL DISORDERS AND ADMINISTRATION SITE CONDITIONS		
Fatigue	410 (46.4%)	46 (52.9%)
Pyrexia	209 (23.6%)	23 (26.4%)
Chills	95 (10.7%)	22 (25.3%)
GASTROINTESTINAL DISORDERS		
Nausea	380 (43.0%)	37 (42.5%)
Constipation	234 (26.5%)	23 (26.4%)
Vomiting	185 (20.9%)	24 (27.6%)
Diarrhea	188 (21.3%)	31 (35.6%)
NERVOUS SYSTEM DISORDERS		
Headache	260 (29.4%)	20 (23.0%)
Dysgeusia	70 (7.9%)	18 (20.7%)
RESPIRATORY, THORACIC, AND MEDIASTINAL DISORDERS		
Epistaxis	223 (25.2%)	22 (25.3%)
Cough	181 (20.5%)	29 (33.3%)
Dyspnea	131 (14.8%)	20 (23.0%)
METABOLISM AND NUTRITION DISORDERS		
Decreased appetite	202 (22.9%)	27 (31.0%)
BLOOD AND LYMPHATIC SYSTEM DISORDERS		
Thrombocytopenia	262 (29.6%)	24 (27.6%)
INVESTIGATIONS		
Aspartate aminotransferase increased	208 (23.5%)	23 (26.4%)
Alanine aminotransferase increased	139 (15.7%)	20 (23.0%)
MUSCULOSKELETAL AND CONNECTIVE TISSUE DISORDERS		
Arthralgia	178 (20.1%)	15 (17.2%)
PSYCHIATRIC DISORDERS		
Insomnia	105 (11.9%)	18 (20.7%)
SKIN AND SUBCUTANEOUS TISSUE DISORDERS		
Rash	115 (13.0%)	19 (21.8%)

[a] Individual AEs with an incidence of > 20% in at least one pooled group are shown.

Appendix 5: Rationale for sensitivity analyses not repeated by ERG

(shaded cells are repeated by the ERG)

	Base case value (BCV)	High value	Low value	Repeated or reason for exclusion
Costs				
PFS supportive care cost (weekly)	£43.45	£86.91 (BCV x2)	£21.73 (BCV x 0.5)	Repeated
Progressed disease supportive care cost (weekly)	£63.08	£123.99 (BCV x2)	£31.54 (BCV x 0.5)	Believe the base case to be an underestimate, so only tested doubling
Post-progression lines (2L/3L) of treatment duration	4.3 months (18.63 weeks)	8.6 months (37.27 weeks)	2.15 months (9.32 weeks)	Does not affect relative costs/ QALYs
Treatment dose	Planned treatment dose	Actual treatment dose observed in the trial for all whole duration of progression-free survival		Base case is actual dose; SA is planned dose
Drug costs	Including wastage	Excluding wastage (full vial sharing) for all drugs (except vinorelbine as generic and assumed to be made up through compounders)		Repeated
Outcomes (Results from the MTC)				
PFS HR: Tra+Cap vs T-DM1	0.68	0.50 (95% CI)	0.91 (95% CI)	The upper CrIs are repeated with the ERG's MTC values. The lower CrIs would not affect the ICER for T-DM1.
PFS HR: Cap vs T-DM1	0.39	0.29 (95% CI)	0.55 (95% CI)	
OS HR: Tra+Cap vs T-DM1	0.68	0.46 (95% CI)	0.98 (95% CI)	
OS HR: Cap vs T-DM1	0.55	0.41 (95% CI)	0.75 (95% CI)	
PFS HR: Tra+Cap vs T-DM1	0.68	0.54 (excluding CEREBEL and Martin *et al.*)		The ERG's random effects MTC takes into account heterogeneity between trials.
PFS HR: Cap vs T-DM1	0.39	0.35 (excluding CEREBEL and Martin *et al.*)		
OS HR: Tra+Cap vs T-DM1	0.68	0.58 (excluding CEREBEL and Martin *et al.*)		
OS HR: Cap vs T-DM1	0.55	0.52 (excluding CEREBEL and Martin *et al.*)		
Progression-free utility	Different values in arms	0.74 (same values as Lap and Cap arm in all arms)		Repeated
Progression-free utility T-DM1	0.78	0.71 (T-DM1 PFS utilty from TH3RESA trial)		Repeated
Progressed utility	0.50	0.70 (BCV +0.2)	0.30 (BCV -0.2)	Believe the base case to be an underestimate, so only tested +0.2

	Base case value (BCV)	High value	Low value	Repeated or reason for exclusion
Progression free utility T-DM1	0.78	0.98 (BCV +0.2)	0.58 (BCV -0.2)	Increasing or decreasing each of these substantially individually is implausible (above analysis is more clinically reasonable)
Progression free utility Lap + cap	0.74	0.94 (BCV +0.2)	0.54 (BCV -0.2)	
Progression free utility Tra + cap/Tra + vin	0.73	0.93 (BCV +0.2)	0.53 (BCV -0.2)	
Progression free utility Cap/Vin	0.72	0.92 (BCV +0.2)	0.52 (BCV -0.2)	
Parametric functions				
PFS	Kaplan-Meier data with log-normal tail	Kaplan-Meier data with other parametric tails (1) Weibull (2) Exponential (3) Log-logistic (4) Gamma (5) Piecewise exponential tail (one piece)		Only the Weibull parametric tail is thought to be clinically plausible and theoretically appropriate. Also tested Weibull distribution for entire curve
OS	Gamma distribution	Other parametric distributions (1) Weibull (2) Log-logistic (3) Log-normal (4) KM data with piecewise exponential tail (one piece) (5) KM data with piecewise exponential tail (two pieces)		Only the Weibull distribution is thought to be clinically plausible and theoretically appropriate. Also tested Weibull distribution for entire curve
Other				
Cost discount rate	3.5%	6%	0%	Varying both at once was thought to be sufficient (see below)
Health outcomes discount rate	3.5%	6%	0%	
Health and cost discount rates	3.5% both arms	6% both arms	0% both arms	Repeated
Time horizon	10	15	5	Repeated (base case is 15 years)