

University of Sheffield



Trastuzumab for the Treatment of
Primary Breast Cancer in HER2
Positive Women:
A Single Technology Appraisal

May 2006

**Sue Ward
Hazel Pilgrim
Danny Hind**

Acknowledgements

Rob Coleman (Professor of Medical Oncology, University of Sheffield), Michael Dixon (Consultant Surgeon and Senior Lecturer, Edinburgh Breast Unit), Frances Lambert (Macmillan Breast Care Nurse Specialist), David Thomson (Lead Pharmacist, British Oncology Pharmacists Association), Andy Hanby (Professor of Breast Pathology, Leeds Teaching Hospital NHS Trust), Catherine Dickens (Consultant Cardiologist, Leeds Teaching Hospital NHS Trust), Laurence O'Toole (Consultant Cardiologist, Sheffield Teaching Hospitals NHS Trust) and Tim Perren (Consultant Oncologist, Leeds Teaching Hospitals NHS Trust) provided advice during the project. Tim Perren, Myfanwy Lloyd-Jones (Senior Research Fellow, University of Sheffield), Eva Kaltenthaler (Managing Director, SchARR Technology Assessment Group), Jim Chilcott, (Technical Director, SchARR Technology Assessment Group), and Zoe Garrett (Technical Lead, NICE), commented on a draft version of the report. The authors wish to thank all of the above. Responsibility for the accuracy of the report lies entirely with the authors. The authors also wish to thank Andrea Shippam for her help in preparing and formatting the report.

This report was commissioned by the NHS R&D HTA Programme on behalf of the National Institute for Health and Clinical Excellence. The views expressed in this report are those of the authors and not necessarily those of the NHS R&D HTA Programme or the National Institute for Health and Clinical Excellence. The final report and any errors remain the responsibility of the University of Sheffield. Jim Chilcott and Eva Kaltenthaler are guarantors.

Contents

1	SUMMARY	5
1.1	Scope of the submission	5
1.2	Summary of submitted clinical effectiveness evidence	6
1.3	Summary of submitted cost effectiveness evidence	7
1.4	Commentary on the robustness of submitted evidence	8
1.5	Key issues	9
2	BACKGROUND	11
2.1	Critique of manufacturer's description of underlying health problem	11
2.2	Critique of manufacturer's overview of current service provision	11
3	Critique of manufacturer's definition of decision problem	12
3.1	Population	12
3.2	Intervention	13
3.3	Comparators	15
3.4	Outcomes	17
3.5	Time frame	19
4	CLINICAL EFFECTIVENESS	21
4.1	Critique of manufacturer's approach	21
4.2	Summary of Results	27
4.3	Critique of submitted evidence syntheses	34
5	ECONOMIC EVALUATION	37
5.1	Overview of manufacturer's economic evaluation	37
5.2	Critique of approach used	54
5.3	Results included in manufacturer's submission	54
5.4	Validity of results presented with reference to methodology used	56
5.5	Summary of uncertainties and issues	56
6	Additional analysis requested from manufacturer with manufacturer's response & ERG's comments	58
7	Additional work undertaken by the ERG	86
7.1	Clinical effectiveness	86
7.2	Cost-effectiveness	92
8	Discussion	104
8.1	Summary of clinical effectiveness issues	104
8.2	Summary of cost effectiveness issues	105
8.3	Implications for research	106
9	Appendices	107
	Appendix A: Further detail of questions for Roche	107
	Appendix B: Potential Errors and/or Inconsistencies in the Model	110
	Appendix C: Quality Assessment Checklist based on original submission	116

List of Tables

Table 1	Identified studies	22
Table 2	Overall survival	28
Table 3	Disease-free survival	28
Table 4	Breast cancer recurrence (local, distant or contralateral events)	29
Table 5	Local recurrence	29
Table 6	Distant recurrence	30
Table 7	Contralateral breast cancer	30
Table 8	Cardiac events	32
Table 9	Adherence	32
Table 10:	Disease-free survival (comparator arm)	40
Table 11:	Overall survival (comparator arm)	41
Table 12	Discounted base case disaggregated cost-effectiveness results	55
Table 13:	Revised adjuvant trastuzumab drug costs	69
Table 14:	Revised metastatic trastuzumab drug costs	70
Table 15:	Revised trastuzumab administration costs	71
Table 16:	Revised drug and drug administration costs	72
Table 17:	Roche's sensitivity analysis around the transition probability from DFS to metastatic	74
Table 18:	Roche's sensitivity analysis around extrapolation of benefits of trastuzumab	76
Table 19:	Roche's sensitivity analysis of the confidence intervals used in the extrapolation of the trastuzumab arm	77
Table 20:	Reduction in the increase of benefits during the first five years following trastuzumab provision	95
Table 21:	ERG's further sensitivity analysis around extrapolation of benefits of trastuzumab	96
Table 22:	Time between local/contralateral recurrence and metastatic recurrence equivalent for both arms	96
Table 23:	Decrease in overall survival	97
Table 24:	ERG's sensitivity analysis around adjustment transition probability due to time in DFS	99
Table 25:	Examples of research papers giving natural history data for HER2 + women	108
Table 26:	Potential Inconsistencies in the model calculations	111

List of Figures

Figure 1:	Roche model structure (see Roche submission)	38
Figure 2:	HER2 testing	48
Figure 5:	Overall survival	87
Figure 6:	Disease-free survival	87
Figure 7:	Distant recurrence.....	88
Figure 8:	Cardiac toxicity.....	88
Figure 9:	52 weeks' versus 9 weeks' trastuzumab	91
Figure 10:	Sensitivity analysis around rate of recurrence over time in DFS	99
Figure 11:	ERG's base case - cost-effectiveness acceptability curve.....	101
Figure 12:	ERG's revised assumption - cost-effectiveness acceptability curve assuming cardiac adverse events as with anthracyclines	102
Figure 13:	Impact of time horizon on incremental cost per QALY gained	103

List of Abbreviations

DDFS	Distant Disease Free Survival
EBCTCG	Early Breast Cancer Trialists' Collaborative Group
EQ-5D	EuroQol questionnaire Five Dimension scale
ERG	Evidence Review Group
FEC-6	Six cycles of Fluorouracil, Epirubicin and Cyclophosphamide
FISH	Fluorescence <i>In Situ</i> Hybridization
HER2	Human EGF-like Receptor No. 2
HERA	HERCEPTIN Adjuvant
HRQoL	Health Related Quality of Life
ICER	Incremental Cost-Effectiveness Ratio
IHC	Immunohistochemistry
LVEF	Left Ventricular Ejection Fraction
LYG	Life Year Gained
MEDTAP	Medical Technology Assessment and Policy
MUGA	Multiple Gated Acquisition scan
NCRI	National Cancer Research Institute
QALY	Quality Adjusted Life Year
RFM	Reduction For Metastases
RFS	Recurrence Free Survival

1 SUMMARY

1.1 Scope of the submission

The manufacturer's scope restricts the population to women who have had HER-2 positive breast cancer removed by surgery and who have completed adjuvant chemotherapy with any standard cytotoxic regimen. The Evidence Review Group (ERG) note that: (1) HER-2 positive breast cancer is a potentially elastic category, depending on the diagnostic tests used; (2) for safety reasons, it is important that anthracycline containing regimens are not given concurrently, although this should not preclude their being given after trastuzumab; and, (3) the exclusion of those at elevated risk of cardiac events from the clinical trials, has screening cost and capacity implications for the NHS as it tries to reproduce those safe conditions.

The manufacturer's scope restricts the intervention to intravenous trastuzumab given for one year after surgery and after the completion of standard adjuvant chemotherapy. The ERG note that: (1) the treatment duration definition is restrictive by comparison with traditional Department of Health/ NICE remits and 'legitimises' the exclusion of an important trial with a shorter treatment duration from the manufacturer's submission; (2) 11% of women in the manufacturer's pivotal trial received neo-adjuvant, not adjuvant, chemotherapy.

The manufacturer's scope restricts the comparator to "Standard therapy without trastuzumab", by implication, NICE's recommended 4-8 cycles of anthracycline-containing chemotherapy post-surgery and 5 years' hormonal therapy. The ERG note that: (1) there is considerable heterogeneity in underlying anthracycline-containing regimens, both within the NHS and the relevant clinical trials; and, (2) it is unclear whether taxanes are to become part of "standard chemotherapy".

The manufacturer's scope defines the primary outcome as disease-free survival (cancer recurrence or death from any cause); secondary outcomes include overall survival, breast cancer recurrence and cardiotoxicity.

Economic outcomes include Cost per Life Year Gained (LYG) and Cost per Quality Adjusted Life Year (QALY) gained. The ERG note that: (1) time is an important factor in breast cancer, which has a long natural history, with recurrences occurring out beyond 20 years; (2) the median follow-up in the pivotal trial is only one year; (3) many consider disease-free survival a surrogate for long-term, all-cause mortality in breast cancer; this has only been empirically demonstrated in other classes of treatment (standard cytotoxics and tamoxifen); and, (4) the manufacturer reasons that the empirically known, short-term harm-benefit profile of trastuzumab will result in a long-term harm-benefit profile similar to that empirically known for other classes of drug.

The ERG also note that: (1) outcome data used in the economic model do not exist in the public domain; (2) the manufacturer has undertaken additional individual patient data analyses to derive these outcomes; (3) the manufacturer have been unable or unwilling to provide the ERG with that dataset so that their work can be checked.

1.2 Summary of submitted clinical effectiveness evidence

The published evidence reports that 18 three-weekly cycles of trastuzumab produced a relative reduction in the hazard of all-cause mortality from 24% (HR 0.76, 95% CI 0.47 to 1.23; absolute risk reduction 0.5%), at a median follow-up of one year in the HERA trial, to 33% (HR 0.67, 95% CI 0.48 to 0.93; absolute risk reduction 1.8%), at a median follow-up of two years in the combined B-31 & N9831 analysis. When all studies with available data were meta-analysed there was a 30% relative improvement in overall survival and this was statistically significant at the five percent level (HR 0.70, 95% CI 0.53 to 0.92, $p=0.010$).

The excluded study, which looked at nine weekly cycles of trastuzumab produced a relative reduction of the hazard of all-cause mortality of 59% (HR 0.41, 95% CI 0.16, 1.08; absolute risk reduction 6.9%) at a median follow-up of three years (the longest follow-up available for any trastuzumab schedule). This study had a small sample size and was not statistically significant at the five percent level.

All studies, at whatever schedule or length of follow-up showed a statistically significant difference in the risk of recurrence or death from any cause (disease-free survival) favouring trastuzumab. The combined hazard ratio for 18 three-weekly cycles was 0.50 (95% CI 0.44-0.57, $p < 0.00001$). In the study evaluating nine weekly cycles the hazard ratio was 0.42 (95% CI 0.21 to 0.83, $p = 0.01$).

There was a statistically significant (almost six-fold) increase in the relative risk (5.54 95% CI 2.07 to 14.82, $p = 0.0007$) of a serious, life-threatening or fatal cardiac event in women treated with 18 three-weekly cycles of trastuzumab, although this represents an absolute risk increase of just 1.6%. In the study evaluating nine weekly cycles there was no excess toxicity.

1.3 Summary of submitted cost effectiveness evidence

Roche have developed a state transition cohort model to compare the lifetime impact of one year of adjuvant trastuzumab therapy to no trastuzumab following standard chemotherapy. The main data source for the model is the HERCEPTIN Adjuvant (HERA) trial¹, an international, multi-centre, randomised trial on women with HER2 positive primary breast cancer, with a median of one year follow-up. Outcomes from the HERA trial are extrapolated over a lifetime horizon in order to assess the long-term benefits and costs of trastuzumab. The model takes into account cardiac toxicity, but does not consider other adverse events. The health states used within the model are considered to be appropriate for the required analysis.

The cost of trastuzumab has been underestimated in the Roche submission, along with the cost of monitoring for cardiac toxicity. The cost and utility associated with each health state were based upon studies carried out by the MEDTAP (Medical Technology Assessment and Policy) research centre specifically for the model. These costs appear high relative to other recent breast cancer models.^{2,3}

The Roche model estimated that the base-case incremental cost-effectiveness ratio (ICER) of chemotherapy plus trastuzumab versus chemotherapy is £5687 per QALY gained, rising to a maximum of £8689 upon

one-way sensitivity analysis of the parameters. However, in the view of the ERG several of the baseline costs were underestimated and some of the upper or lower parameter values tested within the sensitivity analysis were not sufficiently extreme. In addition, there was no sensitivity analysis around the extrapolation of rate of recurrence in the comparator arm and limited sensitivity analysis around the relative risk of recurrence for trastuzumab. With respect to the probabilistic sensitivity analysis, the description of uncertainty surrounding the mean values of many of the model parameters is insufficient or incomplete.

However, following responses from Roche to queries raised by the Evidence Review Group (ERG) in a letter dated 8th March 2006 a revised basecase of £2,387 was presented by Roche (this report, Section 6). Based on additional sensitivity analysis carried out by the ERG (this report, Section 7), the ERG conclude that although the ICER presented by Roche is considered to be too low, the ICER is not expected to rise above £35,000 - £50,000.

1.4 Commentary on the robustness of submitted evidence

1.4.1 Strengths

- The model structure is appropriate and allows sensitivity analysis to be carried out easily.
- One-way sensitivity analysis suggests that variations in the majority of the parameters do not have a large effect upon the ICER.
- The baseline ICER is relatively modest, such that potential parameter variations are unlikely to increase the ICER beyond the currently accepted threshold values.

1.4.2 Weaknesses

- No sensitivity analysis has been undertaken to explore the impact of uncertainty surrounding the comparator arm on the ICER.
- Little sensitivity analysis has been carried out around the long-term benefits of trastuzumab.

- Confidence intervals of some of the parameters do not adequately describe the uncertainty. For instance, the upper values of the cost of trastuzumab and cardiac monitoring were considered to be unrealistic.

1.4.3 Areas of uncertainty

- Disease-free and overall survival may differ from the comparator arm in the model, depending on the chemotherapy regimens being used in the UK.
- The benefits of trastuzumab on rates of recurrence are unknown beyond three to four years.
- There is little evidence to date of the effects of trastuzumab upon overall survival.
- There is no evidence of the effects of trastuzumab upon long term cardiac dysfunction.

1.5 Key issues

The following issues have the potential to impact on the cost effectiveness results. The combined effects of these uncertainties has the potential to increase the ICER from below £5,000 to around £20,000 to £30,000

- The uncertainty generated by long term extrapolation of the comparator arm.
- The uncertainty surrounding the long term benefits of trastuzumab in terms of reduction in the risk of recurrence.
- The extent to which reductions in the rate of recurrence will translate into benefits in overall survival.
- The extent to which patients in both the comparator arm and the trastuzumab arm are likely to receive trastuzumab in the metastatic setting

The addition of potential long term cardiac events could push the ICER above £30,000 although there is no long term evidence to date surrounding this issue.

There are also a number of other important issues which are not explicitly taken into account in the economic modelling

- A small study (the FINHER trial,⁴ n=229), excluded from the manufacturer's submission, raises the possibility of an equally effective but shorter regimen, incurring lower cost and toxicity but with greater patient convenience.
- Capacity issues: HER2 testing, the preparation and administration of trastuzumab and cardiac monitoring will all require the augmentation of currently available facilities.

2 BACKGROUND

2.1 Critique of manufacturer's description of underlying health problem

Roche have provided a sufficient description of general breast cancer, followed by details of Human EGF-like Receptor No. 2 (HER2) positive tumours. More information would have been useful around the average length of disease-free survival for HER2 positive women.

2.2 Critique of manufacturer's overview of current service provision

Current service provision has been briefly outlined, although little detail has been provided as to the current standard chemotherapy treatment in the UK. The comparator arm of the model is based on the HERCEPTIN Adjuvant (HERA) trial, a multi-centre European study (two relevant treatment arms: n=3,387), in which a variety of chemotherapy regimens were used and 26% of patients were on taxanes as well as anthracycline chemotherapy. The HERA trial suggests that the chemotherapy regimen administered does not affect the relative recurrence rates for patients on the trastuzumab arm. However, it should be noted that in the combined analysis of the US trials, in which 100% of patients were on taxanes, the baseline risk of recurrence in the comparator arm was lower. This baseline risk will impact on the cost effectiveness ratio.

3 CRITIQUE OF MANUFACTURERS DEFINITION OF PROBLEM

3.1 Population

The manufacturer's submission defines the population as: "Women with HER-2 positive breast cancer, confined to the breast and draining lymph nodes (early breast cancer) that has been surgically excised and who have completed adjuvant chemotherapy with any standard cytotoxic regimen." Two points of clarification are required: (1) the nature of standard treatment (this report, Section 3.1.1); and, (2) the eligibility of women at high risk of cardiac failure (this report, Section 3.1.2).

3.1.1 Standard treatment

The following clause is from the population definition in the manufacturer's submission (Section 1.1, paragraph 2): '...who have completed adjuvant chemotherapy with any standard cytotoxic regimen'. Clarification is required concerning two issues: (1) the definition of "standard cytotoxic regimen"; and, (2) the role of taxanes.

On the first point, the ERG infers from this statement that, in line with the National Cancer Research Institute (NCRI) guidance,⁵ the manufacturer precludes the use of trastuzumab except in those who have already completed an anthracycline-containing regimen; that is, it should not be given concurrently with anthracyclines. The draft summary of product characteristics in the annex of the submission appears to support this: "Herceptin and anthracyclines should not be used currently in combination except in a well-controlled clinical trial setting with cardiac monitoring". This is in line with the NCRI recommendations (NCRI Guidance, Section 4.4⁵).

On the second point, the manufacturer should make it clear that 'any standard cytotoxic regimen' does not currently include taxanes. However, it should be made clear a patient eligible for a course of trastuzumab might receive it concurrently with a course of taxanes (NCRI Guidance, Section 4.1⁵).

3.1.2 Other criteria for treatment in the NCRI guidance

Note that the NCRI (Guidance, Section 3.1⁵). also require that women:

- have normal left ventricular ejection fraction (LVEF) (as measured on echocardiography or multiple gated acquisition [MUGA] scanning in accordance with local protocol);
- have none of the following cardiac contraindications:
 - a history of documented congestive heart failure;
 - myocardial infarction (unless very good long term);
 - prognosis confirmed by cardiologist);
 - uncontrolled hypertension;
 - unstable arrhythmias; and,
- have an adequate baseline hepatic, renal and haematological function.

The draft updated Summary of Product Characteristics (SPC) in the annex of the report limits these factors to ‘warnings’, rather than ‘contraindications’ (see also manufacturer’s submission, page 13). However, the protocols of the four included trials^{6,7,8,9} preclude participation of women who do not have these characteristics. The trials, then, have no external validity (generalisability: see also this report, Section 4.3.2) with regard to the population contraindicated by the NCRI guidelines.⁵

3.2 Intervention

The manufacturer’s submission defines the intervention as, “Trastuzumab (Herceptin) administered intravenously (IV) for one year after surgery for HER-2 positive early breast cancer and completion of standard adjuvant chemotherapy and radiotherapy if applicable.” The following examines how the definition can be considered appropriate (this report, Section 3.2.1), why the definition of treatment duration might be considered unusually restrictive (this report, Section 3.2.2), and the questionable basis for the proposed restriction of the review to that treatment duration.

3.2.1 The sense in which the definition is ‘correct’

It could be argued that the manufacturer’s definition of the intervention is *de facto* appropriate, because: (1) there is no formal scoping exercise in NICE’s

Single Technology Appraisal (STA) process and the definition of the intervention is supposed to be derived from the license; (2) at the time of writing, the Committee for Medicinal Products for Human Use (CMPHU) has issued no license; and, (3) the CMPHU's license will be based on a submission by the manufacturer. In effect, the manufacturer has been allowed to define the intervention element of the scope and study exclusions which follow from their criteria cannot be disputed on the grounds of internal validity.

3.2.2 Why the definition is restrictive

It could also be argued that the intervention definition is inappropriate because it creates inclusion criteria for the duration of treatment which are unusually narrow compared with the remits given to NICE by the Department of Health for technology assessment reports. The latter are rarely (if ever) prescriptive about dosage or length of treatment. Whether or not the stipulation that trastuzumab should be given for one year is maintained in the marketing authorisation remains to be seen. The matter is of particular interest in this technology appraisal because the manufacturer's scope enables the exclusion of a study which evaluates a shorter treatment schedule (see this report, Section 4.1.3), however reduced treatment duration "may facilitate lower cost, greater patient convenience, and reduced risk of cardiotoxicity."¹⁰

3.2.3 The clinical rationale for the treatment duration

The manufacturer's submission makes no attempt to explain or justify the treatment duration of one year, but the key publication from the HERA trial¹ (which is the basis for the manufacturer's application to the CMPHU) does so in three ways (citations are as given in the HERA publication¹):

- (1) a major peak in the rate of relapse occurs 18 to 24 months after surgery;¹¹
- (2) effective treatment of HER2-positive breast cancer may require prolonged attenuation of HER2 activity;¹²
- (3) tamoxifen, which is an effective targeted therapy for breast cancer, is most beneficial when given for longer than one year.¹³

The elevated risk of breast cancer recurrence during the second year after surgery is well documented.¹⁴ However, it is worth stating that (with the exception of comparatively rare new primary tumours) so-called ‘recurrences’ or ‘relapses’ are actually late clinical expressions of micro-metastases (metastases that had been too small to be identified by conventional means) which have actually been there since the time of diagnosis. In other words, the disease is not absent then present again; it is there all along. There is no necessary causal relationship between length of treatment and disease expression at 18 to 24 months, which is, rather, a function of natural history when the disease is not eradicated in the first place.

The authors’ proposition that treatment requires prolonged attenuation of HER2 activity is based on one laboratory study on xenografts (a surgical graft of tissue from one species to an unlike species). The use of the verb, ‘may’, indicates that this is a possibility rather than a fact.

Finally, the authors’ statement that, “tamoxifen... is most beneficial when given for longer than one year”,¹ whilst not incorrect, is a peculiar reading of the Oxford Overview’s actual findings that, “5 years is significantly... more effective than just 1-2 years of tamoxifen”.¹³ Importantly, the extent to which the analogy with a hormone antagonist (tamoxifen) *is* appropriate is also open to question. The British National Formulary classifies trastuzumab as a cytotoxic drug and the current treatment durations of traditional cytotoxics rarely exceed six months.

3.2.4 Conclusion

In summary, the definition of the intervention as specified by the manufacturers is in accordance with the proposed marketing authorisation. The clinical rationale for the duration of treatment is open to question and leads to the exclusion of one relevant trial.

3.3 Comparators

The Manufacturer’s submission (Section 3; page 7/81) defines the comparator as, “standard adjuvant therapy without trastuzumab”, later adding that NICE currently recommends four to eight cycles of anthracycline-based

chemotherapy for those at intermediate or high risk of recurrence.¹⁵ This is reasonable insofar as gene amplification and overexpression of HER2 is associated with elevated risk of recurrence.^{16,17} Two points of clarification are needed on the character of standard adjuvant therapy: (1) variety in anthracycline-containing regimens currently used in the NHS (this report, Section 3.3.1); and, (2) whether taxanes will be part of baseline treatment (this report, Section 3.3.2).

3.3.1 Anthracycline-containing chemotherapies

It is known that there is a wide variety of anthracycline-containing chemotherapies in use in the NHS, but there is no reliable data on how many and with what frequency each is used. Another NICE Single Technology Appraisal determined, through industry-led market-research, registry data and trial publications, that six cycles of fluorouracil, epirubicin and cyclophosphamide (FEC-6) was the most common regimen and was used by around half of all women with early breast cancer who were eligible for chemotherapy in the UK (including women with HER2 positive tumours).¹⁸ Assuming that there are differences in treatment effect between regimens there are also consequences for external validity (generalisability) if they are used with different frequencies in the included trials.

Differences in baseline comparability with respect to anthracycline containing regimens received would also have important consequences for internal validity (the strength of the evidence) if there were significant differences in their clinical effectiveness. None of the included studies describe the baseline comparability between arms of anthracycline-containing regimens used. All included studies state that they stratified by chemotherapy type but it is unclear whether this is limited to the presence or absence of *any* anthracycline and *any* taxane. In stratification (which reduces the chance of imbalance due to important prognostic variables) there are pragmatic limits to the number of factors, and it is likely that the variety of (differentially effective) anthracycline containing regimens will have been simplified to prevent an unwarranted administrative burden.

3.3.2 Taxanes

Subsequent to the commissioning of this report, a NICE Appraisal Committee has discussed the use of taxanes (in addition to anthracycline-based chemotherapies) for the adjuvant treatment of patients with operable node-positive breast cancer. The final outcome of that appraisal process is still pending, but it follows that if taxanes were to become part of standard treatment, there would be consequences for the external validity (generalisability) of the manufacturer's submission. It is notable that, in the manufacturer's submission, a proportion of participants received taxanes in every included study.

3.4 Outcomes

3.4.1 Lack of clarity in the manufacturer's submission

Roche outline the results of the HERA trial, but do not describe relevant primary and secondary end-points of the model. The primary end-point of the economic model is disease-free survival, with the secondary end-points being overall survival, relapse-free survival and distant disease-free survival. These are appropriate in order to assess the effectiveness of trastuzumab. The definitions of the model outcomes (minus that for 'Reduction For Metastases') are instead described in the manufacturer's submission (page 39, Section 2.3.4) as "described in the HERA study protocol" and it is unclear whether the same definitions apply to the model (for further discussion, see the current report, Section 6, Question A2).

Similarly, the description of the decision problem does not outline possible adverse events. It does suggest that the major adverse event resulting from trastuzumab administration is cardiotoxicity, however there is no description of the way in which this outcome is modelled. The submission does not provide sufficient detail of the way in which health related quality of life (HRQoL) was measured and valued within the model.

3.4.2 Choice of primary endpoint

The manufacturer's primary outcome was disease-free survival, a composite of disease recurrence, contralateral breast cancer, death without prior

recurrence and other second primary cancer. Composite endpoints are appropriate when a single primary variable cannot be selected from multiple measurements associated with the primary objective.¹⁹ However, statistical efficiency is a more common motivation in time-to-event trials, where composite outcomes lead to higher event rates and thus enable smaller sample sizes, shorter follow-up or both.²⁰ It is for this reason that disease-free survival is a popular choice as a primary outcome in evaluating interventions for early stage cancer, although it is a choice that is not without its critics. In the words of a leading British breast surgeon:

“in clinical science, there are only two meaningful outcome measures — length and quality of life. All other outcome measures are surrogate, and that includes disease-free survival.”²¹

Apologists argue that, where long-term follow-up is available (for instance, in the Oxford Overview¹³), there is a correlation between short-term disease-free survival and long-term all-cause mortality, but a correlate does not necessarily make a reliable surrogate.²² The expectation that disease-free survival *automatically* translates into overall survival or quality of life is erroneous, because unexpected adverse events may increase deaths from other causes and decrease quality of life in the novel treatment arm, thus obliterating the benefit in cause-specific deaths.²³ In short, whilst disease-free survival may be related to overall survival, it is unlikely that this relationship is directly proportional.

3.4.3 Transparency of outcomes

Unlike relative risks and odds ratios which can be re-calculated when event numbers are reported, time-to-event outcomes are rather more opaque because they factor in: (1) the time at which an event happened; and, (2) the censoring of patients who will never be recorded as having an event because of loss to follow-up or a ‘competing risk’ (such as death, where an outcome does not include death as an event). Without access to individual patient outcomes data, the ERG has to take summary statistics which are published in peer review journals on trust. But, where outcomes not available in the public domain are calculated for some specific purpose (for instance

economic modelling), additional caution is urged. This point is revisited in the current report, Section 6 (in our comment on the manufacturer's response to Question A1) and throughout.

3.5 Time frame

Although the risk of breast cancer recurrence is highest in the second year, it may still be substantial during the second decade. Meta-analyses¹³ suggest that, compared with no chemotherapy, most of the effect of polychemotherapy on the risk of recurrence is seen in the first five years after randomisation, although that benefit is maintained thereafter. However, in terms of breast cancer mortality, while there is some gain from polychemotherapy during the first five years, the absolute gain is generally at least twice as great for 15-year survival as for five-year survival. It follows that a five-year study is not long enough to demonstrate the full impact of adjuvant chemotherapy on overall survival.

It is the case that breast cancer which over-expresses HER2 is likely to recur earlier¹⁶ and, on that basis, it is likely that health gains will become apparent more quickly. This is an important consideration in the interpretation of three of the included studies which were stopped early. HERA was stopped at the interim analysis, with a median follow-up of one year (maximum follow-up of three years) because improvements in disease-free survival were highly significant ($p < 0.0001$) and satisfied a pre-specified stopping rule.¹ Data from the NSABP B-31 and parts of the NCCTG N-9831 trials were pooled for a joint interim analysis of results (median follow-up two years; maximum follow-up 4.5 years) when both were stopped for the same reason.²⁴ No one in any of the included trials has been followed up for five years, the most commonly-used proxy time-point for long-term survival in the solid tumours.

Two questions arise when studies are stopped early on the grounds of benefit: (1) whether currently significant clinical benefits would still have been significant at a later time point; (2) whether currently non-significant harms would be significant by a later time point.^{25,26,27} These issues are discussed further below (this report, Section 4.1.7).

The manufacturer's health economic model has used a lifetime horizon of 45 years to assess the long-term benefits of trastuzumab, which is justified. However, owing to the short duration of follow-up (median of one year) of the HERA trial, numerous important assumptions were made in order to extrapolate outcomes for both treatment arms over a lifetime horizon. The long term benefits of trastuzumab are, therefore, subject to considerable uncertainty. Inadequate sensitivity analysis around these assumptions was presented in the manufacturer's submission.

4 CLINICAL EFFECTIVENESS

4.1 *Critique of manufacturer's approach*

4.1.1 Description of manufacturer's search strategy

The manufacturer's searches, conducted in December 2005, were restricted to publications from 1993 onwards. Whether and how restrictions were placed on the searches (as opposed to the study selection, for which, see this report, Section 4.1.2) is not clear. The searches were simple and relied heavily on MeSH headings without support from free-text terms. There was also inconsistency between searches of different databases. For instance, some used 'trastuzumab' and 'herceptin' as search term, whilst others just used 'trastuzumab'. Nevertheless, the search strategies were adequate to retrieve important citations relating to all eligible studies of which the ERG and its clinical advisors are aware.

4.1.2 Statement of the inclusion/exclusion criteria used

Restrictions to studies on the adjuvant use of trastuzumab (and synonyms) in humans with early breast cancer and (given the timescale) English language publications were appropriate.

Restriction to 'clinical trial data publications' (presumably meaning '*controlled* clinical trials' – those with comparator arms) is appropriate for the assessment of clinical benefit. However, the reporting of clinical harms is often inadequate in controlled clinical trial publications because they exclude patients at high risk from harms,²⁸ may be too short to identify long-term or delayed harms, or may have sample sizes too small to detect uncommon events.^{29,30,31,32}

As discussed in Section 3.2 of this report, the definition of the intervention under review was proscriptive, and the consequences are discussed in the next section.

4.1.3 Identified and excluded studies

4.1.3.1 Identified studies

The manufacturer's submission (page 27; Section 2.2.1; Table 1) identified the five relevant phase III clinical trials of which the ERG are aware: HERA (n=3,387),¹ BCIRG-006 (n=2,148),³³ NCCTG N9831 (n=1,615), NSABP B-31 (n=1,736)²⁴ and FinHER (n=229).⁴ The study characteristics are re-tabulated in simplified form (Table 1).

Table 1: Identified studies

Trial	Population	Underlying interventions	Trastuzumab schedule
HERA (Europe) n=3,387	Premenopausal 16%; Postmenopausal 47%; Uncertain 38%; Node negative 33%; Tumour <2cm 40%; ER+ve 45%	Neoadjuvant chemotherapy: 11%; No anthracyclines 6%; Anthracyclines no taxanes 68%; Anthracyclines and taxanes 26%;	Loading dose 8 mg/kg then 6 mg/kg 3-weekly for 1 or/2 years.
NCCTG N9831 (USA) n=1,615	<50 yrs: 50%; >50 yrs 50%; Node negative 12%; Tumour <2cm 39%; ER+ve 52%.	Neoadjuvant chemotherapy: 0%; Anthracyclines 100%; Taxanes 100%;	Loading dose 4mg/kg then 2 mg/kg weekly for 51 weeks
NSABP B-31 (USA) n=1,736	<50 yrs: 52%; >50 yrs 48%; Node negative 0%; Tumour <2cm 39%; ER+ve 52%.	Neoadjuvant chemotherapy: 0%; Anthracyclines 100%; Taxanes 100%;	Loading dose 4mg/kg then 2 mg/kg weekly for 51 weeks
B-31 & N9831 (Combined analysis) n=3,351	<50 yrs: 51%; >50 yrs 49%; Node negative 6%; Tumour <2cm 39%; ER+ve 52%.	Neoadjuvant chemotherapy: 0%; Anthracyclines 100%; Taxanes 100%;	Loading dose 4mg/kg then 2 mg/kg weekly for 51 weeks
BCIRG 006 (USA) n=2,148	<50 yrs: 52%; >50 yrs: 48%; Node negative: 29%; Tumour <2cm 40%; HR+ve: 54%	Neoadjuvant chemotherapy: 0%; Anthracyclines 100%; Taxanes 100%;	Loading dose 4m/kg then 2 mg/kg weekly during taxane-delivery then 6mg/kg 3-weekly for total of one year
FinHer (Finland) n=229	<50 yrs: 50%; >50 yrs: 50%; Node negative: 16%; Tumour <2cm 35%; ER+ve: 55%	Neoadjuvant chemotherapy: 0%; Anthracyclines 100%; Taxanes 48%	Loading dose 4 mg/kg then 2mg/kg weekly for nine weeks.

4.1.3.2 Excluded studies

The manufacturer's submission (page 28; Section 2.2.2) excludes FinHER from the review on two grounds:

1. Neither of the two underlying chemotherapy regimens "would be much used in the UK".
2. Trastuzumab "was administered according to a dose schedule very different from that used in other studies and this trial cannot be considered a test of the intervention under review".

In FinHER, trastuzumab was given weekly for nine weeks, concurrently with three cycles of docetaxel or eight cycles of vinorelbine, *before* three cycles of FEC. With this in mind, the manufacturer's first point (that the underlying anthracycline-containing regimen is not generalisable to the UK) has merit on two grounds: (1) in terms of quantity, NICE currently recommends four to eight cycles of anthracycline-based chemotherapy for those at intermediate or high risk of recurrence¹⁵ (see this report, Section 3.3); and, (2) in terms of timing, the NCRI (Guidance,⁵ page 6, paragraph 4.1), preclude the use of trastuzumab except in those who have already received an anthracycline-containing regimen (although, importantly they explicitly based their statement on the published evidence which, at the time, did not include data from FinHer).

However, exclusion on these grounds is contrary to the logic of Roche's repeated statement (Manufacturer's submission, pages 4, 21, 50 and 62) that: "regardless of...type of breast cancer surgery or adjuvant chemotherapy, trastuzumab reduces the risk of a DFS event to a similar extent." It also makes inconsistent the inclusion of the American trials which, by Roche's own admission contain important differences in underlying treatment strategies (manufacturer's submission, page 27, Section 2.2.2). Furthermore, the same logic brings into question how good a basis for generalisation the HERA trial provides: only 68% of its study population received adjuvant anthracyclines without taxanes (see Table 1, above).

The second point (nine weeks' trastuzumab "cannot be considered a test of the intervention under review") is technically correct because, as discussed in Section 3.2 of this report, the manufacturer has been allowed to define the intervention under review and has done so in an unusually narrow fashion, limiting inclusion to studies which evaluated 52 weeks of trastuzumab. Variations in the trastuzumab's length of use could then usefully have been highlighted along with variation in concomitant therapies in Section 1.3, part 24 of the manufacturer's submission: "Describe any current variation in services and/or uncertainty about best practice, including cost effectiveness".

Finally, we should note that, prior to the ERG's receipt of the manufacturer's submission, one of our oncologist expert witnesses expressed a belief that the manufacturer should not exclude the FinHer study from their report. But, he also expressed the opinion that the FinHer's value lay mainly in hypothesis generating rather than problem solving.

In summary, we believe the manufacturer should have included this study, although it is not without problems. We address the predictive value of FinHer, with its implications for cost, patient convenience, and safety, in Section 7.1 of this report.

4.1.4 Details of any relevant studies that were not included in the submission

Between the 9th and 28th March, 2006, the ERG re-ran the manufacturer's searches. No further Phase III controlled trials were found when the search results were screened.

4.1.5 Description and critique of manufacturer's approach to validity assessment

The manufacturer's approach to validity assessment was adequate. However, it is worth noting some confusion in the writing of NICE's report template, which could be improved on two issues: (1) the difference between randomisation and allocation concealment; and, (2) the definition of intention-to-treat analysis.

4.1.5.1 Randomisation and allocation concealment

In Section 2.4.6 of the template, question 46 asks, "Which of the following best describes the randomisation?" The choices offered refer to concealment of the allocation schedule, not the generation of the sequence itself, as implied in the question. Inadequate concealment of the allocation schedule has been empirically demonstrated to exaggerate clinical effect sizes more than any other factor,³⁴ but it is also important to validate that the method of randomisation is truly random. The current template does not allow the manufacturer to demonstrate this.

4.1.5.2 Intention-to-treat analysis

In Section 2.4.2 of the template, question 47 asks, “Which of the following best describes the adequacy of follow-up?” The final choice is clearly supposed to be adequate: “(C) Trial outcome(s) were assessed in all treated and control subjects.” However, a generous reading of this, could be applied to a per protocol analysis, where only those who are treated are analysed. The statement could better read, “...in all *randomised* subjects in the treatment and control arm” to ensure that it only reflects an intention-to-treat analysis (the most reliable kind³⁴).

4.1.6 Description and critique of manufacturer’s outcome selection

The manufacturer has taken a structured approach to the reporting of disease-free survival (see this report, Section 3.4, for a critique of the primary outcome) and overall survival, but a more selective and non-systematic approach to “other secondary efficacy parameters” (manufacturer’s submission, page 54). There is no explanation given for the omission of local recurrence and contralateral disease as outcomes, although these are commonplace in evaluations of interventions for early breast cancer. This is not an important point, as event rates are comparatively small for both outcomes and, as such, have less bearing on overall survival than distant recurrence.

4.1.7 Describe and critique the statistical approach used

The manufacturer’s submission reproduces summary statistics from trial publications, in full for disease-free and overall survival, but more selectively for “other secondary efficacy parameters”. For instance, confidence intervals are missing in Table 14 (manufacturer’s submission, page 54), which also reproduces redundant data on overall and disease-free survival. In systematic reviews it is customary to tabulate results from studies side by side, which the manufacturers have failed to do in outcomes other than disease-free survival.

The most uncomfortable aspect of the presentation of outcome data in both the HERA study paper (Figure 2)¹ and the manufacturer’s submission (page 48, Table 13) is the presentation of event rates and absolute risk differences

for two years follow-up. Only about 15 percent of the randomised participants had been followed up for this length of time and, although the trial paper (if not the manufacturer's submission) presents confidence intervals for these estimates, these too should be interpreted with caution. It is not unknown for highly significant treatment effects, identified in such small sample sizes after a short follow-up, to be completely reversed in instances where data monitoring committees allow trials to continue.³⁵ Whilst such a reversal *seems* unlikely in the HERA trial, its early closure ensures that, by itself, it will never *directly* address the issue of clinical utility in overall survival or provide a proper long-term harm-benefit analysis in the unique population it randomised.

The weight of the evidence from other, more mature studies suggests that the early benefits of trastuzumab *should* be maintained, at least over the first four years after surgery (the maximum length of time for which any reasonable numbers have been followed up^{4,24,33}). By analogy with other treatments for early breast cancer, most would expect these effects to be maintained over the long-term.¹³ Nevertheless, the presentation of statistics exaggerates what is known from the HERA sample population in terms of what is, in any case, a surrogate outcome (disease-free survival – see this report, Section 3.4).

4.1.8 Summary statement

4.1.8.1 Completeness of the submission: studies and data

As far as the ERG are aware, the manufacturer's submission is complete with regard to studies that evaluate 52 weeks of trastuzumab. The most important data from within those studies are present, but there is inconsistency, redundancy and presentation bias (see this report, Section 4.1.7) in reporting.

4.1.8.2 Does the submitted evidence reflect the decision problem?

Short-term follow-up and the early termination of three studies mean that the HERA data is inadequate to *directly* address the issue of clinical utility in overall survival or provide a proper long-term harm-benefit analysis. The selective use of data from other studies, none of which have run for five years, does little to rectify this problem. This aside, the submitted evidence adequately reflects the narrow decision problem as defined in the submission

(see this report, Section 3.2 for commentary). A broader perspective is presented in Section 7.1 of the current report.

4.2 Summary of Results

Results of all identified trials abstracted from the study publications by the ERG are tabulated below. Absolute Risk Reductions (ARR) and Numbers Needed to Treat (NNT) have been estimated by the ERG where possible.

4.2.1 Overall survival

Where reported, the relative reduction in the hazard of all-cause mortality ranged from 24% at a median follow-up of one year in the HERA study (0.5% absolute hazard reduction) to 59% at a median follow-up of three years in the FinHer study (6.9% absolute hazard reduction: see Table 2). Only the B-31 & N9831 combined analysis (median follow-up two years) found a statistically significant difference between trastuzumab and observation, with a hazard reduction of 33% (1.8% absolute hazard reduction). It is worth noting that clinical trials of cancer treatments used in the adjuvant setting are rarely adequately powered to detect significant differences in overall survival and, when these are apparent, it is usually after longer follow-up.

4.2.2 Disease-free survival

Every study demonstrated a significant reduction in the hazard of recurrence or death (disease-free survival). Absolute hazard reductions ranged from five percent in N9831 to ten percent in B-31 (Table 3).

4.2.3 Breast cancer recurrence

Local and distant recurrences are combined with contralateral events in Table 4, to give a composite estimate of breast cancer recurrence; the three component outcomes are reported separately in Tables 5, 6 and 7 respectively.

Table 2: Overall survival

Trial	Follow-up	N followed up		Mortality						
	(months)	Trastuzumab	Ctrl	Trastuzumab		Ctrl		HR (95% CI)	ARR (95% CI)	NNT
				n	%	n	%			
HERA	12	1694	1693	29	1.7	37	2.2	0.76 (0.47,1.23)	0.005 (-0.008,0.018)	211.17 (56.81,∞)
B-31 & N9831	24	1679	1672	62	3.7	92	5.5	0.67 (0.48,0.93)	0.018 (-0.002,0.038)	55.26 (26.08,∞)
B-31	29	864	872	NR	NR	NR	NR	Not reported	Not estimable	Not estimable
N9831	19	808	807	NR	NR	NR	NR	Not reported	Not estimable	Not estimable
BCIRG 006	23	1075	1073	20	1.9	36	3.4	Not reported	0.015 (not estimable)	66.91 (not estimable)
FinHer	36	115	116	6	5.2	14	12.1	0.41 (0.16,1.08)	0.069 (not estimable)	14.43 (9.97,-111.10)

Table 3: Disease-free survival

Trial	Follow-up	N followed up		Recurrence or Mortality						
	(months)	Trastuzumab	Ctrl	Trastuzumab		Ctrl		HR (95% CI)	ARR (95% CI)	NNT
				n	%	n	%			
HERA	12	1694	1693	127	7.5	220	13.0	0.54 (0.43,0.67)	0.055 (0.027,0.083)	18.19 (12.09,36.76)
B-31 & N9831	24	1679	1672	133	7.9	261	15.6	0.48 (0.39,0.59)	0.077 (0.046,1.078)	13.01 (9.28,21.76)
B-31	29	864	872	83	9.6	171	19.6	Not estimable	0.100 (0.067,0.133)	10.00 (7.52,14.89)
N9831	19	808	807	50	6.2	90	11.2	Not estimable	0.050 (0.022,0.076)	20.14 (12.99,44.85)
BCIRG 006	23	1075	1073	77	7.2	147	13.7	0.49 (0.37,0.65)	0.065 (not estimable)	15.30 (not estimable)
FinHer	36	115	116	12	10.4	27	23.3	0.42 (0.21,0.83)		

Table 4: Breast cancer recurrence (local, distant or contralateral events)

Trial	Follow-up	N followed up		Recurrence						
	(months)	Trastuzumab	Ctrl	Trastuzumab		Ctrl		HR (95% CI)	ARR (95% CI)	NNT
				n	%	n	%			
HERA	12	1694	1693	118	7.0	220	13.0	Not reported	0.060 (not estimable)	16.59 (not estimable)
B-31 & N9831	24	1672	1679	126	7.5	256	15.2	Not reported	0.077 (0.046,0.1080)	12.96 (9.26,21.63)
B-31	29	864	872	77	8.9	152	17.4	Not reported	0.085 (not estimable)	11.74 (not estimable)
N9831	19	808	807	43	5.3	85	10.5	Not reported	0.052 (not estimable)	19.19 (not estimable)
BCIRG 006	23	1075	1073	NR	NR	NR	NR	Not reported	Not estimable	Not estimable
FinHer	36	115	116	11	9.6	26	22.4	0.46 (not reported)	0.128 (not estimable)	7.782 (not estimable)

Table 5: Local recurrence

Trial	Follow-up	N followed up		Loco-regional recurrence						
	(months)	Trastuzumab	Ctrl	Trastuzumab		Ctrl		HR (95% CI)	ARR (95% CI)	NNT
				n	%	n	%			
HERA	12	1694	1693	27	1.6	59	3.5	Not reported	0.019 (not estimable)	52.88 (not estimable)
B-31 & N9831	24	1672	1679	27	1.6	57	3.4	Not reported	0.018 (not estimable)	56.18 (not estimable)
B-31	29	864	872	15	1.7	35	4.0	Not reported	0.023 (not estimable)	43.90 (not estimable)
N9831	19	808	807	12	1.5	22	2.7	Not reported	0.012 (not estimable)	80.58 (not estimable)
BCIRG 006	23	1075	1073	NR	NR	NR	NR	Not reported	Not estimable	Not estimable
FinHer	36	115	116	2	1.7	6	5.2	Not reported	0.034 (not estimable)	29.13 (not estimable)

Table 6: Distant recurrence

Trial	Follow-up	N followed up		Distant recurrence						
	(months)	Trastuzumab	Ctrl	Trastuzumab		Ctrl		HR (95% CI)	ARR (95% CI)	NNT
				n	%	n	%			
HERA	12	1694	1693	85	5.0	154	9.1	0.49 (0.38,0.63)	0.041 (0.017,0.065)	24.52 (15.49,58.76)
B-31 & N9831	24	1679	1672	96	5.7	193	11.5	0.47 (0.37,0.61)	0.058 (not estimable)	17.17 (not estimable)
B-31	29	864	872	60	6.9	111	12.7	Not reported	0.058 (not estimable)	17.29 (not estimable)
N9831	19	808	807	30	3.7	63	7.8	Not reported	0.041 (not estimable)	24.43 (not estimable)
BCIRG 006	23	1075	1073	NR	NR	NR	NR	Not reported	Not estimable	Not estimable
FinHer	36	115	116	8	7.0	26	22.4	0.29 (0.13,0.64)	0.15 (not estimable)	6.59 (5.22,13.47)

Table 7: Contralateral breast cancer

Trial	Follow-up	N followed up		Contralateral						
	(months)	Trastuzumab	Ctrl	Trastuzumab		Ctrl		HR (95% CI)	ARR (95% CI)	NNT
				n	%	n	%			
HERA	12	1694	1693	6	0.4	7	0.4	Not reported	0.000 (not estimable)	1,687 (not estimable)
B-31 & N9831	24	1672	1679	3	0.2	6	0.4	Not reported	0.002 (not estimable)	592.02 (not estimable)
B-31	29	864	872	2	0.2	6	0.7	Not reported	0.005 (not estimable)	219.01 (not estimable)
N9831	19	808	807	1	0.1	0	0.0	Not reported	-0.001 (not estimable)	-808 (not estimable)
BCIRG 006	23	1075	1073	NR	NR	NR	NR	Not reported	Not estimable	Not estimable
FinHer	36	115	116	3	2.6	1	0.9	Not reported	-0.017 (not estimable)"	-57.25 (not estimable)

4.2.4 Adverse events: cardiovascular events

In what follows, we comment on the quality of the data presentation in the manufacturer's submission (Section 4.2.4.1), re-present data in a more consistent and transparent format (Section 4.2.4.2) and respond to the manufacturer's perspective on trastuzumab toxicity (Section 4.2.4.3).

4.2.4.1 Quality of data presentation in the submission

The manufacturer presents outcome data in Table 16; however outcomes are ambiguously-defined in Table 15 of their submission, making the data essentially useless.

- It is not clear what is being counted in Table 16. Raw event numbers are not given, only percentages, which are not easily associable with data in study publications. We are referred to Table 15 for definitions but are not told whether the bulleted endpoints in the second column of Table 15 were added together as a composite endpoint or whether only one is being counted.
- It is not clear at what time point events are being counted (as with other endpoints, time is likely to be a factor with heart damage). Column three in Table 15 mentions short-term safety analyses, but it would be more appropriate to report all events reported at the time of the primary efficacy analysis, as in the study publications.
- If the latter was the case, there would be no need for the (unexplained) confidence intervals in Table 16, the presence of which suggest a modelled (rather than observed) estimate of effect.

4.2.4.2 Re-presentation of data

Table 8 re-presents data as abstracted by the ERG from the available trial publications. We grouped together events that were described as serious, life-threatening or fatal, as defined by the different classificatory systems used in the trials.^{36,37}

Table 8: Cardiac events

Trial	Follow-up	N followed up		Cardiac Events						
	(months)	Trastuzumab	Ctrl	Trastuzumab		Ctrl		HR (95% CI)	ARR (95% CI)	NNT (Harm)
				N	%	n	%			
HERA†	12	1694	1693	9	0.5	2	0.1	NR	-0.004 (-0.008,-0.0002)	242 (125,3736)
B-31 & N9831†	24	1679	1672	31	1.8	5	0.3	NR	-0.015 (-0.023,-0.008)	65 (44,119)
BCIRG 006‡	23	1075	1073	25	2.3	10	0.9	NR	-0.014 (-0.025,-0.003)	72 (40,320)
FinHer	36	115	116	0	0.0	NR	NR	NR	Not estimable	Not estimable

† New York Heart Association class III or IV congestive heart failure or death from cardiac causes; ‡ Common Toxicity Criteria for Adverse Events Grade 3-4 cardiac left ventricular function (CHF), cardiac ischemia/infarction or arrhythmias

Table 9: Adherence

Trial	Follow-up (months)	N followed up	Trastuzumab withdrawal	
			N	%
HERA*	12	1694	143	8.4
B-31 & N9831†	24	1679	337	20.1
BCIRG 006	23	1075	NR	NR
FinHer	36	115	NR	NR

The absolute risk of a serious, life-threatening or fatal cardiac event increased by between 0.4% (the HERA trial¹ after a median follow-up of one year) and 1.5% (the combined B-31 & N9831 analysis²⁴ after a median follow-up of two years).

When symptomatic congestive heart failure was added to these numbers, the risk difference increases considerably. Symptomatic and asymptomatic decreases in left ventricular ejection fraction account for the majority of withdrawals from trastuzumab in all trials. In the HERA trial, 5.5% of women randomised to trastuzumab withdrew from treatment before completion due to adverse events (8.4% withdrew overall – see Table 9). A higher rate of withdrawals in the combined US study is explained by the manufacturer in terms of a difference in withdrawal criteria:

“The NSABP B-31/NCCTG N9831 studies used different criteria for withdrawal for asymptomatic LVEF decline (14.2% for this reason alone in the NSABP B-31/NCCTG N9831 studies versus 6.0% for all adverse events in HERA). These may have led to the withdrawal of patients who would have continued on treatment and asymptomatic in HERA” (Manufacturer’s submission, page 57, Section 2.8). As the HERA trial forms the basis of Roche’s trastuzumab submission, the HERA withdrawal rate is more likely to be representative.

4.2.4.3 Cardiotoxicity in context

Most people would agree with the manufacturer that, compared with the hazard of recurrence, the risk of a serious, life-threatening or cardiac event, at least in the short-term, is low. They are correct that, by externalising the cost of screening and monitoring to the health care system, so that the target population more closely resembles that of the trial populations, serious cardiac problems can be minimised (manufacturer’s submission, page 60).

The manufacturer cites papers which characterise trastuzumab cardiac damage as essentially short-term and reparable (if one survives the acute decline) in contrast to standard cytotoxic therapies.^{38,39,40} The late cardiotoxic effects of anthracyclines are known to become apparent between four and 24

years after initial therapy.⁴¹ The cardiologists interviewed by the ERG, while acknowledging that trastuzumab affects the heart in a different way from anthracyclines, characterised this work as speculative. Their experience was that almost all patients with dilated cardiomyopathy (global weakness of heart muscle, as opposed to coronary artery disease) respond well, initially, to standard medical therapy. However, they also expressed a concern that this was driving a degree of undue optimism: many people with dilated cardiomyopathy will present decades later with end-stage heart failure or sudden death. At the same time, there are acute cardiomyopathies from which people appear to recover completely. Into which category trastuzumab will eventually fit can only be ascertained by long-term follow-up and, until that is available, cardiologists advise monitoring according to the HERA trial protocol. Whilst this may prove overly cautious, the patients involved are predominantly relatively young women, most of whom can expect to live for decades.

It is worth restating that all of this refers to potential trastuzumab damage *over and above* that from which women are already at risk from anthracyclines and, where available, taxanes.

4.3 Critique of submitted evidence syntheses

4.3.1 The strength of the evidence (internal validity)

The search strategy was poorly designed (this report, Section 4.1.1) but the ERG have not determined that any relevant primary studies were missed as a result (this report, Section 4.1.4). The inclusion criteria were adequately defined (this report, Section 4.1.2), but the manufacturer's study selection and use of the published evidence seemed to work on a highly selective and arbitrary basis, in the reporting of outcome data (this report, Sections 4.1.3, 4.1.6 and 4.1.7). The manufacturer's approach to validity assessment appears to have been adequate, but the template they were asked to use by NICE has problems (this report, Section 4.1.5). The manufacturer's reporting of secondary outcomes, particularly adverse events was somewhat haphazard (this report, Section 4.2.4.1).

Critical outcomes used in the model were poorly defined in the manufacturer's submission and not reported in the public domain. As the review team could not access the individual patient data from the pivotal trial, they are unable to validate the manufacturer's analysis of this data. Where comparisons with similar published outcomes are possible there is no evidence of any inexplicable discrepancies (this report, Section 3.4.1).

Time is an important factor in breast cancer, which has a long natural history, with recurrences occurring out beyond 20 years. The median follow-up in the pivotal trial is only one year. Many consider disease-free survival a surrogate for long-term, all-cause mortality in breast cancer. This has only been empirically demonstrated in other classes of treatment (standard cytotoxics and tamoxifen). The manufacturer reasons, by analogy alone, that the empirically known, short-term harm-benefit profile of trastuzumab will result in a long-term harm-benefit profile similar to that empirically known for other classes of drug (this report, Sections 3.5, 4.1.7 and 4.1.8.2).

4.3.2 The applicability of the results (external validity)

Women at elevated risk of a cardiac event were not recruited to the clinical trials which evaluated trastuzumab (this report, Sections 3.1.2 and 4.3.2) . Those women who were recruited were intensively monitored. This puts the onus (and the additional cost of screening) on the NHS to replicate an eligible population for whom the treatment will be as safe as in the clinical trials. If the current shortfall in cardiac monitoring capacity is not adequately addressed, women treated with trastuzumab will be at elevated risk of heart failure compared with those in the clinical trials.

A restrictive scope allowed the manufacturer to exclude from any serious discussion the FinHer study (Section 4.1.3.2). The manufacturer rightly pointed out that the underlying anthracycline-containing regimen was different to any used in the NHS. However, cancer clinicians have noted that the nine week regimen examined in this study may facilitate lower cost, greater patient convenience, and reduced risk of cardiotoxicity, although the evidence is not as strong as that for 52 weeks.

Considerable heterogeneity of study populations in terms of the concomitant chemotherapies received and lack of knowledge about what regimens are in use in the NHS make generalisation from the published evidence problematic (Section 3.3.1, 3.3.2), although the direction and extent of clinical effect seems relatively consistent across different baseline treatment programmes.

4.3.3 Summary

The published evidence reports that 18 three-weekly cycles of trastuzumab produced a relative reduction in the hazard of all-cause mortality from 24% (HR 0.76, 95% CI 0.47 to 1.23; absolute risk reduction 0.5%) at a median follow-up of one year to 33% (HR 0.67, 95% CI 0.48 to 0.93; absolute risk reduction 1.8%) at a median follow-up of two years.

The excluded study, which looked at nine weekly cycles of trastuzumab produced a relative reduction of the hazard of all-cause mortality of 59% (HR 0.41, 95% CI 0.16,1.08; absolute risk reduction 6.9%) at a median follow-up of three years (the longest follow-up available for any trastuzumab schedule). This study had a small sample size and was not statistically significant at the five percent level.

All studies, at whatever schedule or length of follow-up, showed a statistically significant relative difference (of about 50% in every case) in the hazard of recurrence or death from any cause (disease-free survival) favouring trastuzumab.

In every study, there was a statistically significant increase in the relative risk of a serious, life-threatening or fatal cardiac event in women treated with 18 three-weekly cycles of trastuzumab, although this represented a small absolute risk increase. In the study evaluating nine weekly cycles there was no excess toxicity.

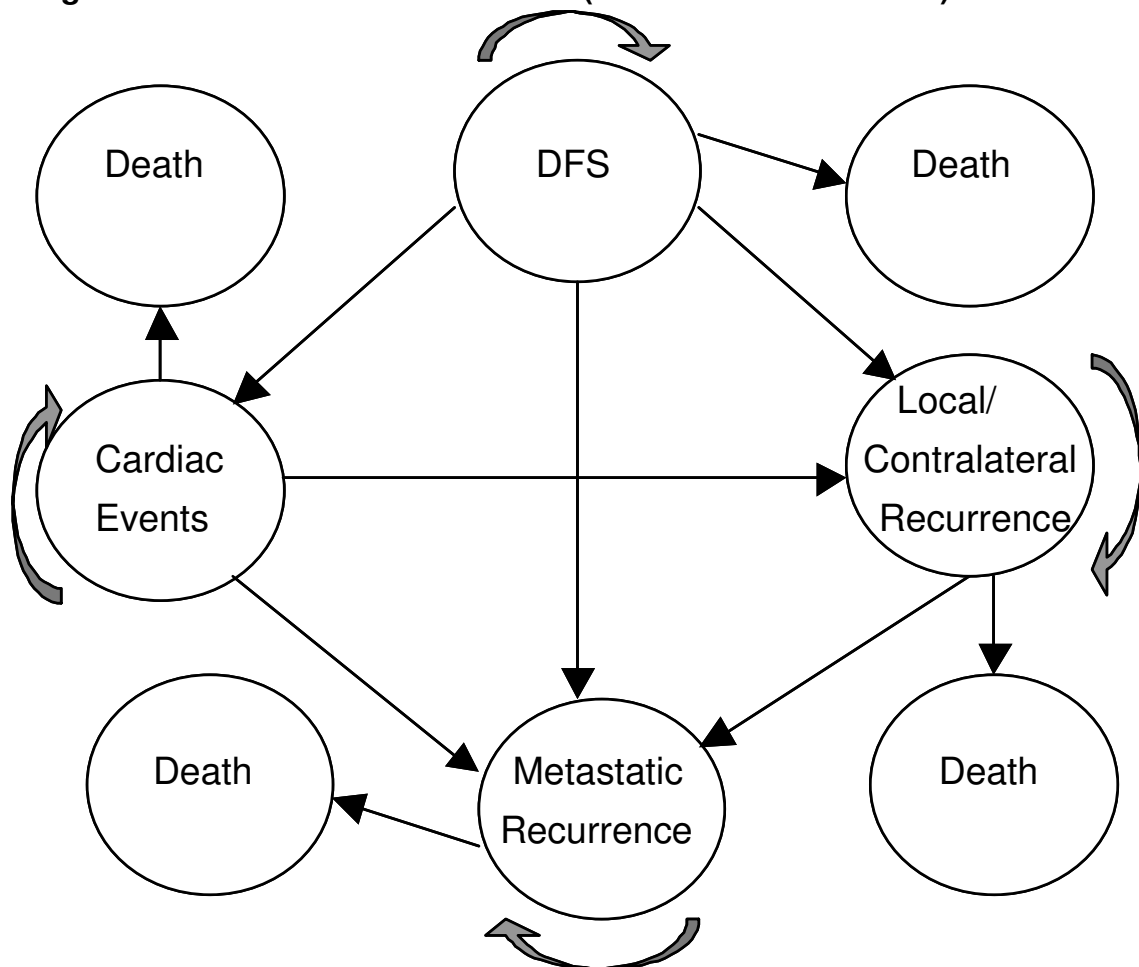
5 ECONOMIC EVALUATION

5.1 Overview of manufacturer's economic evaluation

A state transition cohort model was used to compare the lifetime impact of one year of adjuvant trastuzumab therapy to no trastuzumab following standard chemotherapy regimens based on the HERA trial. The clinical effectiveness aspect of the model is based upon the HERA trial¹ which was an international, multi-centre, randomised trial on women with HER2 positive primary breast cancer. All women had completed locoregional therapy and at least four cycles of neo-adjuvant or adjuvant chemotherapy, following which 1694 women received one year of trastuzumab whilst 1,693 women were allocated to the observation group. In addition, 1,694 women were allocated to a third arm to receive two years of trastuzumab, however these results are not yet available and this analysis has not been included in the model.

The health states included within the model and their relationship to each other are shown in Figure 1 below. Both the costs of each health state and the health utility scores were provided by MEDTAP (in Appendices 2 and 3 respectively of the manufacturer's submission). Transition probabilities were taken from the HERA trial. The exception to this was the probability of moving from the metastatic state to death which was sourced from a previous Roche study of trastuzumab in the metastatic setting.⁴²

Figure 1: Roche model structure (see Roche submission)



The key assumptions of the model are:

- The effect of time on the rate of recurrence is relative to that collected by the Early Breast Cancer Trialists' Collaborative Group (EBCTCG)¹³ based on all women with breast cancer
- The same trastuzumab relative risk of recurrence collected from the HERA trial of **(CIC information removed)** can be applied until year 10 when it is increased by a third (to a relative risk of **(CIC information removed)**) until year 45.
- All-cause mortality is applied to patients in the Disease Free and Recurrence Health State as well as in the Cardiac Adverse Event State.
- Only 1 loco-regional/contralateral breast cancer recurrence can be experienced per patient.

- After a single year in the locoregional/contralateral recurrence state, the patients return to a health state with a similar utility and cost as the DFS state. Patients remain there until they die or experience metastases.
- Immunohistochemistry (IHC) and fluorescence *in situ* hybridization (FISH) testing will be used to determine HER2 status. FISH testing will be carried out on around 10% of patients who receive borderline (2+) IHC results.
- Cardiac monitoring will be carried out at baseline, 3, 6 and 9 months.
- Occurrence of a cardiac event does not increase mortality.

5.1.1 Natural history

The disease natural history of HER2 positive patients was modelled by using transition probabilities from the HERA trial database, which has a median follow-up of one year and a maximum follow-up of thirty-six months. These data were used to extrapolate over forty-five years.

The rate of recurrence for the comparator arm for the first year is taken from the HERA trial. The model assumes that the rate of recurrence varies over time and the trend of recurrence over time is based on the recurrence rates over time of all women with early breast cancer taken from by the Early Breast Cancer Trialists' Collaborative Group (ECGBCG) overview¹³ On this basis, for the first five years recurrence rates are assumed to be the same as in the first year, whereby recurrence was reduced by a factor of 0.64 and 0.41 at years 5 and 10 respectively. Women with HER2 positive breast cancer have a higher risk of recurrence and hence by definition do not have the same disease natural history as other breast cancers. However, advice from our expert clinicians suggest that assuming the same pattern of recurrence over time is a fair assumption given the lack of any other evidence with regard to this subgroup.

There is limited long-term evidence regarding the disease natural history of HER2 positive women; however several studies^{43,44} including the American

trials, NCCTG N9831 and NSABP B-31,²⁴ suggest that the model may have over-predicted the number of recurrences and over-estimated survival in the comparator arm, therefore overestimating the benefits of trastuzumab. Tables 10 and 11 compare the disease-free and overall survival at 2, 5 and 10 years for HER2 positive patients following a number of different chemotherapy regimens taken from research papers.

Table 10: Disease-free survival (comparator arm)

Source	Chemotherapy regime	Probability of disease-free survival (comparator arm)		
		At 2 years	At 5 years	At 10 years
Model data	Various regimens	<i>(CIC information removed)</i>		
American trial data ²⁴	Doxorubicin, cyclophosphamide and paclitaxel	0.84	-	-
FinHer trial data ⁴⁵	Docetaxel or vinorelbine plus CEF and tamoxifen	0.84	-	-
Moliterni et al ⁴⁶	CMF (plus doxorubicin)	0.87 (0.87)	0.62 (0.66)	0.46 (0.57)
Mastro et al ⁴⁷	FEC21	0.83	0.63	0.46 (9yrs)
Stal et al ⁴³	Tamoxifen	0.76	0.63	0.59
Nieto et al ⁴⁸	High-dose chemotherapy	0.73	0.62	0.59
Press et al ⁴⁴	No chemotherapy	0.82	0.72	0.58

NB. Probabilities taken from research papers may have been projected from graphs.

Table 11: Overall survival (comparator arm)

Source	Chemotherapy regime	Probability of overall survival (comparator arm)		
		At 2 years	At 5 years	At 10 years
Model data	Various regimens	<i>(CIC information removed)</i>		
American trial data ²⁴	Doxorubicin, cyclophosphamide and paclitaxel	0.96	-	-
FinHer trial data ⁴⁵	Docetaxel or vinorelbine plus CEF and tamoxifen	0.96	-	-
Moliterni et al ⁴⁶	CMF (plus doxorubicin)	0.93 (0.93)	0.78 (0.78)	0.54 (0.68)
Mastro et al ⁴⁷	FEC21	0.89	0.75	0.72 (9yrs)
Stal et al ⁴³	Tamoxifen	0.88	0.71	0.61
Nieto et al ⁴⁸	High-dose chemotherapy	0.85	0.68	0.61
Press et al ⁴⁴	No chemotherapy	0.97	0.75	0.61

NB. Probabilities taken from research papers may have been projected from graphs.

The comparisons from these papers should be considered as indicative only as the patient groups are not directly comparable with the population in the model. However table 10 suggests that particularly for long-term estimates such as at 10 years, disease-free survival may have been underestimated by in comparison to other trial data. This difference is likely to be amplified over the 45 year time horizon. Table 11 suggests that the long-term overall survival may have been overestimated in the comparator arm in comparison to other trial data. Therefore, the percentage difference between overall survival and disease-free survival is much greater for the model data than from any other available trial data. All of the above papers suggest that the number of recurrences alleviates after around five years. The ERG has carried out additional sensitivity analysis around these issues (this report, Section 7.2).

The probability of transiting from the metastatic state to death was sourced from a previous Roche study of trastuzumab in the metastatic setting⁴² since these data were not available from the HERA trial. The chemotherapy regimen of doxorubicin (or epirubicin) and cyclophosphamide used within the

study is no longer standard UK practice for metastatic breast cancer; hence the probability of death is likely to be decreased. In comparison with the Slamon study of trastuzumab in the metastatic setting, the time spent in metastases is around (CIC information removed) in the model; however the one way sensitivity analysis carried out by Roche suggests that this should not have a big impact upon the estimated cost-effectiveness of trastuzumab.

The use of taxanes for treatment of early breast cancer is currently under review by NICE. As taxanes have the potential to become standard treatment for breast cancer, their effect on the model results should be considered as this would alter the assumed baseline risk of recurrence. The HERA trial suggests that the relative risk of recurrence for trastuzumab is the same regardless of the chemotherapy regimen previously received. However, since the administration of a taxane may reduce the baseline risk of recurrence, the absolute benefit of trastuzumab would be decreased. (This issue is discussed further in Section 7.2.2.3 of this report).

5.1.2 Treatment effectiveness

The model assumes that the relative risk of recurrence for patients on trastuzumab compared with no trastuzumab observed within the HERA trial in the first year is maintained for the 10 years following initial administration of trastuzumab and that two-thirds of this benefit is seen until year 45. There is however no evidence beyond four years, and little evidence beyond two years, of the effectiveness of trastuzumab from the existing trials, although the American trials²⁴ suggest that the benefits of trastuzumab may continue for 3 or 4 years.

There are examples of treatments for early breast cancer which offer benefits, in terms of reduced risk of recurrence well beyond the treatment period. The EBCTCG overview paper¹³ provides an overview of randomised trials of chemotherapy and tamoxifen suggests that 6 months of anthracycline-based chemotherapy for patients with early breast cancer provides benefits in terms of a reduction in the risk of recurrence for around five years. Also the benefits of 5 years treatment of tamoxifen are shown to reduce the risk of recurrence for around 10 years, demonstrating a protective “carry-over” effect for five

years beyond the treatment period. Advice from the ERGs' clinical advisors has suggested that trastuzumab may follow the same pattern. However, this assumption is subject to significant uncertainty and only a limited sensitivity analysis was carried out by the sponsor around these assumptions.

In addition, the Roche model has applied the same confidence intervals for the relative risk of recurrence for years one to 10 and 10 to 45. As the amount of uncertainty in the number of recurrences and deaths increases over time, these confidence intervals should also increase over time in the Roche model to adequately capture this uncertainty.

5.1.3 Health related quality of life (HRQoL)

MEDTAP carried out a study for Roche to determine utility scores for specific model health states. Standard gamble interviews based on health states established by oncologists, breast cancer specialists and psychometric experts were completed by one hundred members of the public. In addition, all participants completed the EQ-5D to assess their own current health. Further details can be found in Appendix 2 of the Roche submission.

Collecting health utility scores from the MEDTAP study rather than using previously established health utility scores enabled scores to be collected which were representative of HER2 positive women. It should be noted, however, that the study was based on a **(CIC information removed)** The participants were not completely representative of the population in that there were **(CIC information removed)** than in the general UK population. In addition, the average **(CIC information removed)** of the participants was **(CIC information removed)**. Despite this, comparing these health utility scores with those of Karnon et al² suggested that they were in line with utility values used in other recent early breast cancer models.

5.1.4 Resources and costs

5.1.4.1 Cost of each health state

Roche funded MEDTAP to carry out a study to determine the costs associated with each health state. (CIC information removed) to seven physicians who indicated (CIC information removed). However, some bias may have been incorporated (CIC information removed). In addition, a small number of experts provided information.

MEDTAP obtained unit costs from NHS Reference Costs, PSSRU Unit costs for health and social care and their own database; all but the latter are recognised as established resources. It is difficult to comment on the reliability of the MEDTAP database and the frequency of its use is unknown.

The modelled costs of each health state appear high in comparison to other research papers for early breast cancer, such as that by Karnon et al²; however clinical experts have suggested that the breakdown of calculated costs appears reasonable. Uncertainty around these costs is not expected to have a significant impact on the cost effectiveness results.

5.1.4.2 Cost of trastuzumab

The exclusion of administration costs.

The cost of trastuzumab administration was not included in the Roche model. The administration of trastuzumab involves a ninety minute infusion; the patient being closely observed from the start of the infusion for six and two hours for the initial and maintenance doses respectively. This requires an outpatient appointment in a chemotherapy suite and is estimated to cost up to £3,000⁴⁹ in total.

The number of vials required per patient.

The number of trastuzumab vials provided for each patient is likely to have been underestimated. It was assumed that the average weight of a woman receiving trastuzumab is 70kg, meaning that they would require three vials of the drug every three weeks. However, in order to receive two vials, the woman would have to weigh less than 50kg, as against only 75kg to receive four vials. Therefore, the percentage of women requiring four vials is likely to have been significantly underestimated in the model. It has been predicted by the ERG that the additional cost would be around £2,400.

Dose banding

Dose banding involves the calculated dose required by the patient to be rounded up or down according to predetermined standard doses. In this way, doses can be prepared in advance, potentially saving pharmaceutical time and wastage. However, trastuzumab cannot be stored for longer than 24 hours following preparation. Therefore, it is anticipated that dose banding will be employed only in a minority of areas and it is not expected to have a significant effect on the cost or effectiveness of trastuzumab.

Number of maintenance doses required.

Roche have assumed that sixteen maintenance doses will be required, although the current recommended guidelines⁵ suggest that seventeen maintenance doses should be administered. Whilst there may be some variability between oncologists, it is anticipated that the majority will follow the UK recommended guidelines. The additional cost of the extra dose is approximately £1,400.

Capacity issues.

The administration of trastuzumab requires the use of a chemotherapy suite for at least two hours per visit. Some chemotherapy suites will have spare capacity and will be able to accommodate the increased workload, but many centres will have a shortage of available chairs in their chemotherapy suites.

In addition, there is likely to be a lack of capacity for reconstitution. This is a difficult area as some pharmacists would recommend that they themselves reconstitute it, while others suggest it could be carried out by nurses by the patient's chair.

There are two potential solutions to these capacity issues:

(1) Additional funding for NHS chemotherapy suites

(2) Use of non-NHS home chemotherapy delivery services such as Healthcare at Home

(1) An increase in the number of chairs required in the chemotherapy suites would be required to accommodate the increase in outpatient appointments in many areas, meaning substantial investment would be required.

(2) There is potential for private homecare providers to provide the treatment at home by the bedside on behalf of the NHS, which, because no V.A.T. is charged upon the drug, may be cost neutral (David Thomson, personal communication), although exact costs of healthcare at home is at yet unknown. Trastuzumab is classified as a medium-risk drug since (i) it requires constant observation during the infusion and for some time afterwards, (ii) there is a risk of allergic reactions and (iii) it involves multiple vials. However, as trastuzumab poses no risk to the safety of the specialist administering the drug, administration at home is a viable option. The ERG's pharmaceutical advisor suggested that home delivery services such as Healthcare at Home are likely to be used in areas where there is insufficient spare capacity in local chemotherapy suites.

Scheduling.

The model assumes that trastuzumab will be administered following the completion of standard chemotherapy regimens. However, the NCRI UK guidelines state that 'published evidence supports two schedules: a. Following a standard anthracycline based regimen; b. With a taxoid following 4 courses of an anthracycline based regimen. Either schedule is currently acceptable

pending comparative data from the NCCTG-N9831 trial which are not yet available.' Similarly, the model assumes that trastuzumab will be given after radiotherapy; however the NCRI UK guidelines suggest that it may be offered concurrently. If trastuzumab was offered alongside chemotherapies or radiotherapy the costs of administration may be reduced.

Implications of the FinHer trial

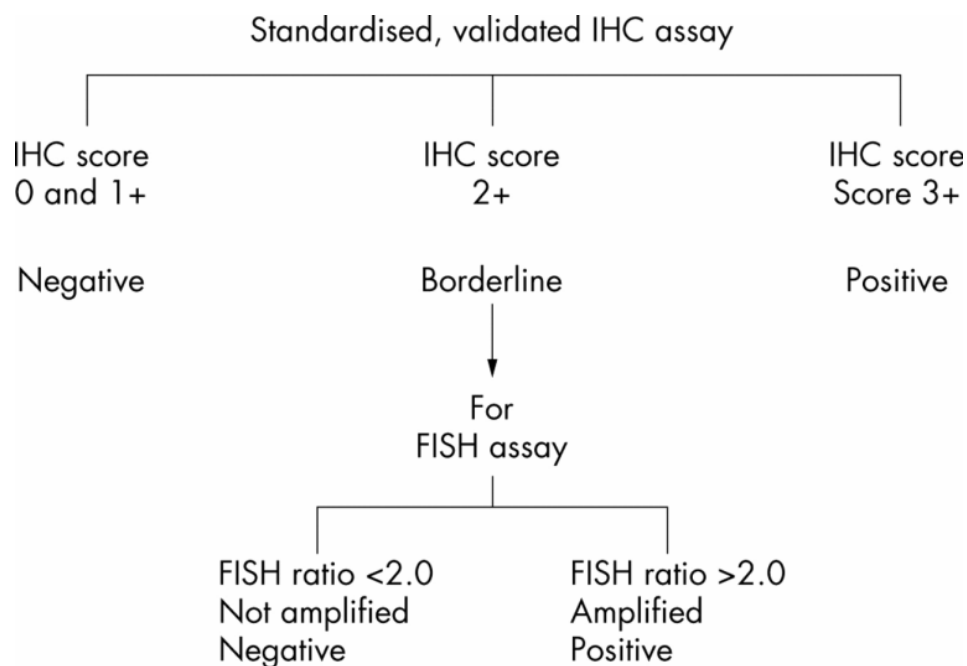
The FinHer trial⁴ involved patients being given nine weekly infusions of trastuzumab instead of an infusion every three weeks for twelve or twenty-four months. The results suggested that trastuzumab is no less effective when given for this shorter period of time, almost halving the costs, although these results should be treated with caution (see section 7.1).

5.1.4.3 HER2 Testing

There are currently two widely used tests for determining whether a patient with breast cancer is HER2 positive: immunohistochemistry (IHC) and fluorescence in situ hybridization (FISH). The IHC test measures the amount of HER2 protein expressed on the surface of tumour cells, whereas the FISH test tests for gene amplification. Some studies have shown it to be more specific and sensitive than IHC.

Patients are defined as HER2 positive if the tests indicate protein overexpression at a 3+ level by IHC or gene amplification by FISH. The majority of UK centres use IHC 3+ as an initial basis for trastuzumab treatment. Tumours scoring 2+ on the IHC test are considered to be borderline and are subsequently tested using the FISH test to determine whether they are HER2 positive (see Figure 2 below). The HERA trial followed this outlined protocol.

Figure 2: HER2 testing



The sensitivity and specificity of the tests.

There are a proportion of tumours which may receive a 3+ IHC score but are FISH negative or may receive a 0/1+ IHC score but are FISH positive. Studies suggest that these percentages could be as great as 6%⁵⁰ and 1.37%⁵¹ respectively. *If the FISH test is considered to be the gold standard*, this could potentially mean that some women tested with IHC would receive trastuzumab without the possibility of benefit (false positives) and some women would not receive the drug despite the fact that they could benefit (false negatives). If as many as 6% of trastuzumab prescriptions related to false positives, the expected cost of provision is increased and the benefits reduced.

Given that the HERA trial used the combined approach of IHC testing, followed by FISH testing for those scoring 2+ with IHC, the impact of false positives or negatives on the cost-effectiveness of trastuzumab will have been encapsulated in the results of the HERA trial and the subsequent economic modelling by Roche. However if more accurate identification of patients who

would respond to trastuzumab treatment was possible it would improve the cost effectiveness of treatment.

The use of FISH testing for all patients has been advocated by some experts. On the basis that this is the gold standard test and that the additional expense of the test would be more than offset by the savings from preventing the treatment of false positives. There is however no clear cut evidence demonstrating that the FISH test is the gold standard in relation to determining whether women will respond to trastuzumab. For example, the HER-2/*neu* Briefing Document from the Food and Drug Administration, 2001²³ suggested that those patients with tumours scoring 3+ on the IHC test may benefit from treatment even if the FISH test result is negative.

Further evidence is needed to determine whether FISH testing offers more accurate identification of patients who will respond to trastuzumab treatment and therefore whether the use of 100% FISH testing offers a more cost effective option for identification of patients eligible for treatment with trastuzumab. This is viewed as an issue for consideration outside the Single Technology Assessment process.

Insufficient staffing and overhead costs appear to have been applied.

The Roche submission identified the cost of IHC and FISH tests as £35 and £70 respectively. These costs are considered to be underestimates and more realistic cost estimates are £50 and £160 respectively⁴⁹. It is not expected that these small increases will significantly affect the model results.

Variation in the percentage of IHC 2+ scores requiring FISH tests.

The percentage of IHC tests scoring 2+ varies greatly between histopathologists, given its subjectivity. On average, around 13% of all samples score IHC 2+⁵², although this could be as high as 30%⁵³. Therefore, applying the costs outlined above and increasing the percentage of FISH tests required, the average cost of HER2 testing may be increased to as much as £95 per test in some cases. However, again, this is not expected to have a significant effect on the model results.

To ensure the highest possible standards for HER2 testing it is recommended that testing is done in recognised laboratories defined by each UK Network. These laboratories should perform at least 250 IHC tests annually. It is also recommended that FISH testing is undertaken in laboratories performing 100 tests or more per year.

Inadequate tests.

An average of around 5% of IHC and FISH tests are required to be repeated due to inadequate samples. This percentage varies between laboratories in the UK from 0% to around 12%. However, this small increase in costs was not expected to have a significant effect upon the ICER.

Role of HER2 testing for other interventions

It was announced in October 2005 that all women with early breast cancer should have HER2 testing. Other treatment decisions may prove to be determined by HER2 status hence it may be considered inappropriate to incorporate the full cost of HER2 testing into the analysis of trastuzumab.

Capacity issues

Additional IHC and FISH testing staff and equipment are likely to be required to meet the increased demand for HER2 testing; therefore, potential further investment may be required in this area.

5.1.4.4 Cardiac Monitoring

Trastuzumab is associated with cardiac adverse events and patients were only eligible for the HERA trial if their LVEF was above 55%. Cardiac monitoring was carried out at baseline, 3, 6, 9 and 12 months using a 2D echocardiogram or MUGA scan. If at any time the patient's LVEF fell below 50% with an absolute reduction of 10% from baseline or decreased to 45%

trastuzumab administration was discontinued. The recent NCRI guidelines suggest a similar protocol.⁵

Cost of monitoring

The model assumes the cost of an ECHO scan to be £120. This may be slightly underestimated. In practice, many centres use the MUGA scan which is a less subjective, more accurate test; however this carries a substantially greater cost of around £258.⁵⁴ If one third of the cardiac monitoring was carried out using MUGA tests, the mean cost would be increased to around £166 per test.

The Roche model assumes that all women will be monitored on four occasions; however the UK clinical guidelines⁵ suggest that women should be assessed five times; at baseline, 3, 6, 9 and 12 months. In addition, Roche have not allowed for the additional appointments required upon a reduction in LVEF. The HERA trial data suggests that an extra cardiac test is likely to be required in around 7% of women. This means that the cost per year could be increased to £888 rather than £475, taking the above increased costs into consideration. There is also a possibility that further monitoring will be required following the provision of trastuzumab.

Accuracy of the scan

A 2D Echocardiogram known as the Simpson's bi-plane method was employed in the HERA trial.¹ A less accurate scan is likely to be used in UK clinical practice. In addition, where 2D echocardiograms are employed rather than MUGA scans, variation in the LVEF scores of around 10-15% is likely to result from the subjectivity of different cardiologists carrying out the tests in practice. As a result, the chances of the women not being prescribed trastuzumab or being taken off the drug unnecessarily may be increased, reducing effectiveness and increasing costs of the drug as a consequence. However, the recommended guidelines⁵ suggest that women should be monitored twice before they are finally taken off the treatment; meaning that an inaccurate decision remains reasonably unlikely.

Currently, there is no evidence regarding the effect of a delay in administration of trastuzumab due to cardiac adverse events upon recurrence rates. However, because of the long half-life of the drug this is expected to have a minor effect upon the relative recurrence rates.

Capacity issues

Many centres in the UK lack the capacity to meet their current demand of echocardiograms, with waiting lists of up to six months. Therefore, providing appointments every three months seems infeasible without substantial extra funds being provided in this area.

Implications of the FinHer trial⁴

Trastuzumab was not associated with cardiac adverse events in the FinHer trial, suggesting that decreases in LVEF and heart failure may be the result of receiving the drug for a longer period of time. If cardiac adverse events could be avoided by providing shorter regimens of trastuzumab, the costs associated with cardiac adverse events and cardiac monitoring would be reduced. In addition, the effectiveness of the drug would be increased since a higher percentage of women would be able to finish the regimen. However, since there were only 232 HER2 positive women participating in the FinHer trial, it would be possible to make a 'type II error' where the hypothesis that cardiac events are more likely to occur in the trastuzumab arm is incorrectly rejected.

Long-term effects

There is currently no evidence regarding the long-term effects of cardiac adverse events as a result of receiving 12 months of trastuzumab. Late cardiotoxic effects of anthracyclines do not become apparent for several years of initial therapy (see section 4.2.4.3) and it is possible that this will also be the case for trastuzumab, although the short term damage is known to be different. The Roche model does not take into account the potential impact of long term effects on the cost-effectiveness of the drug.

5.1.5 Discounting

Roche have assumed a discount rate for both costs and health benefits of 3.5%. This is in line with the current NICE guidance. It should be noted that variations in the discount rate have a significant effect on the model results.

5.1.6 Sensitivity analyses

One-way sensitivity analyses have been carried out on many of the model parameters by Roche; none of which increase the ICER by more than £3,000. However, some of the figures used in the one-way sensitivity analysis were not sufficiently extreme. For example, the upper bound for the cardiac monitoring cost was £713 which may not be sufficiently high to represent the estimated cost of cardiac monitoring discussed in Section 5.1.4.

In addition, several key assumptions were made in order to extrapolate the data over the forty-five years, whilst it appears a limited one-way sensitivity analysis was applied solely to the relative risk of recurrence for the trastuzumab arm over the first five years. Therefore, additional sensitivity analysis should have been carried out surrounding the relative risk from years five to 10, and from 10 to 45. Roche assumed that the relative risk of recurrence taken from the first year of the HERA trial could be applied up to year 10 and then increased by a third from year 10 to 45. Since none of the outlined trials have follow-up of more than four years, and since other trials suggest that the benefits of drugs for high-risk breast cancer diminish after five years, it is considered reasonable to model a scenario whereby there is no benefit in terms of relative risk reduction for patients on trastuzumab after five years.

Furthermore, a comparison of the comparator arm to the other trials outlined in Roche's clinical effectiveness search and to previously published research papers on HER2 positive women suggested that the disease-free survival may have been underestimated in the model and the overall survival may have been overestimated. Examination of the impact of uncertainty was not carried out with regards to the comparator arm in either the one-way or the probabilistic sensitivity analysis. Variation in chemotherapy regimens used, and the potential of taxanes to become the standard chemotherapy in the

near future, should lead to extensive sensitivity analysis to ensure that the model results are robust for all relevant UK women.

Therefore, since insufficient or no uncertainty has been incorporated into the model parameters, Roche's probabilistic sensitivity analysis does not ensure that the model results are robust.

5.1.7 Model validation

Validation was carried out by Roche to ensure that modelled clinical outcomes were representative of the HERA trial; however the Roche submission did not try to validate the assumptions employed surrounding the extrapolation approach. There is limited evidence surrounding the long-term effects of trastuzumab in HER2 positive patients as targeted chemotherapy is a relatively new development. Owing to this uncertainty, more extensive sensitivity analysis should have been carried out around the data extrapolation (see section 5.1.6 above).

In addition, a comparison of the HERA trial data with the data collected from the other identified studies (outlined in section 4.1.3.2, Table 1) would have provided validation of the effectiveness of trastuzumab. However, no meta-analysis was carried out. The ERG have provided their own meta-analysis (see section 7.1).

5.2 Critique of approach used

The state transition model which Roche used is considered to be appropriate for the economic analysis. However, whilst the model is structurally adequate, further validation and sensitivity analysis is required to ensure that the results are robust.

5.3 Results included in manufacturer's submission

The results of the Roche model from the original submission are presented in Table 12 below.

Table 12: Discounted base case disaggregated cost-effectiveness results

	Trastuzumab	No trastuzumab	Incremental
Costs (£)	87159	73323	13835
Life year gained	14.11	11.69	2.43
QALYs	11.21	8.78	2.43
Cost per LYG	£5702		
Cost per QALY	£5687		

The critical appraisal of the Roche model undertaken by the ERG suggests that the base case ICER is optimistic. This is largely due to the underestimation of the cost of trastuzumab and assumptions concerning the ongoing benefits of the drug over the patient's lifetime (see section 5.1).

One-way sensitivity analyses were carried out on the majority of the parameters. Variability in the cost of trastuzumab, the cost of being in the DFS state and the discount rates had the greatest impact upon the ICER. These analyses resulted in ICERs between £8,000 and £9,000.

The probabilistic sensitivity analysis presented within the submission suggested that the probability that trastuzumab has a cost-effectiveness ratio that is better than £20,000 per QALY gained is close to 1. Importantly, the description of uncertainty surrounding several of the model parameters was inadequate; hence the true uncertainty surrounding the incremental costs and benefits of trastuzumab is unlikely to have been captured (see section 5.1.6).

From this point in the present report onwards, Incremental Cost-Effectiveness Ratios (ICERs) refers to the cost per Quality-Adjusted Life Year (QALY) gained rather than the Life Year Gained (LYG).

5.4 Validity of results presented with reference to methodology used

As discussed above, the ICER of £5,687 is likely to be optimistic and insufficient sensitivity analysis and validation was carried out to ensure that the model results are robust.

However, it seems unlikely that the ICER will rise above £35,000 - £50,000 providing that trastuzumab provides a benefit in terms of risk of recurrence similar to that seen in the first year of the HERA trial for at least the first five years. Our clinical advisors have agreed that benefits are likely to be maintained for at least five years based on the evidence presented to date from current trials, such as the NCCTG N9831 and NSABP B-31,²⁴ and the past experience of previous drugs for high-risk breast cancers.

5.5 Summary of uncertainties and issues

The Roche model is reasonably robust in terms of the cost parameters applied. Increasing costs of HER2 testing and cardiac monitoring is expected to have only a marginal effect on the results. Increasing the costs of trastuzumab has the largest impact of any of the costs upon the model results, increasing the ICER by several thousands of pounds.

The amount and length of time for which trastuzumab will reduce recurrence rates may have been over-predicted in the model. Altering the assumptions regarding extrapolation of the data has the greatest impact on the model results. However, evidence from the literature and clinicians suggests that it is likely that the drug will have an effect for at least five years. Using this as a worse-case scenario gives an ICER of less than £20,000. However, combined consideration of all other uncertainties is likely to increase the cost-effectiveness of trastuzumab further.

The main aspects of the Roche report and model which would benefit from improvements are as follows.

Inclusion of the FinHer trial data⁴ in sections 2.2, 2.3 and 2.4 of the report. ⁴

- A sensitivity analysis surrounding the extrapolation of the HERA comparator arm (particularly to assess the effect a taxane being the standard treatment).
- A sensitivity analysis of the future relative recurrence rates in the trastuzumab arm.
- An analysis of potential long term cardiac adverse events.
- Increasing the predicted cost of cardiac monitoring, HER2 testing and trastuzumab.
- Ensuring the upper and lower bounds of each of the parameters are sufficiently large or small to capture the true uncertainty around mean values.
- A discussion around current capacity issues.

6 Additional analysis requested from manufacturer with manufacturer's response & ERG's comments

A letter was sent to Roche on 8th March 2006 from NICE/ERG requesting clarification of various issues in the submission and requesting a range of further analyses. Roche responded with a detailed letter on 28th March 2006. The NICE/ERG's questions are shown in bold, followed by the response from Roche. The ERG's comments upon the manufacturer's responses are shown in the boxes following each question. Greater detail of the questions can be found in Appendix A.

Section A: Clarifications

A1 Please clarify how the trastuzumab risk reduction values used in the model were derived from the HERA trial. It does not seem as though they were drawn from the published paper.

The risk reductions/hazard ratios used in the health economic model for trastuzumab are based upon a separate analysis of the HERA clinical trial database compared to that analysis utilised within published articles. The reason for this supplementary analysis is that the hazard ratios reported in the published papers use composite end-points that are neither applicable nor valid in the context of the economic model structure. Therefore a number of recodings and censorings took place, explained in greater detail in question A3 below in order to ensure transition probabilities were representative of the health states within the economic model.

This explains the correct observation within the question above that the hazard ratios differ from those within the published paper. The hazard ratios are calculated using a standard treatment only adjusted Cox model based upon trial data that are censored using the algorithm explained in question A3 below.

A composite endpoint is a clinical outcome which includes more than one type of event. The primary outcome of the HERA trial is disease-free survival (DFS), which incorporates any cancer recurrence or death without recurrence.¹ The manufacturer is correct that the composite endpoints used in clinical trial papers, whilst of use to clinicians, are rarely ideal for economic modellers. There is nothing unusual about these outcomes not being in the public domain (especially 'reduction from metastases': A2, below), but the fact that they are only accessible to the manufacturer by a fresh analysis of individual patient data underlines the lack of transparency in the current process. There is no means by which the ERG can work backwards from the equations in Section A3 and check the manufacturer's work.

The **Cox proportional hazards regression analysis**⁵⁵ is widely used in survival analysis when investigators wish to investigate a number of prognostic factors in addition to treatment effect. In the industry submission, the manufacturer has only used the Cox model in order to derive hazard ratios for the model. They have not adjusted for prognostic variables ("treatment only adjusted Cox model"), which is legitimate because baseline comparability is satisfactory in the HERA trial.¹ Hazard ratios can be considered like relative risks, although they take into account *when*, as well as *whether*, an event occurs: it is the "relative survival experience" of the two treatment groups.⁵⁶

A2 Please define the acronyms RFS, DDFS and RFM used in the model. These outcomes did not appear to be published in the HERA trial paper (Piccart-Gebhart 2005) and while some definition is given in the SABCS abstract, more detailed information would help our interpretation of these outcomes. Please also give definitions of these outcomes used in the ASCO abstract.

RFS

The term RFS refers to "Recurrence Free Survival", within the economic model this is a variable used to adjust the baseline transition probabilities for those patients who received adjuvant trastuzumab. In the approved protocol BO16348D, dated 13 November 2003, RFS is defined as the time from randomization to the date of first local, regional or distant tumour cancer (section 8.1.2.2). However within the context of the economic model, the RFS

variable is the hazard ratio for developing a recurrence, where recurrence is defined as a local manifestation of the old (primary) cancer e.g. cancer cells left on the scar after mastectomy. The model includes as a separate event contra-lateral breast cancer. The hazard ratio for this event is (CIC data removed) but has overlapping confidence intervals with 1. Therefore, in the economic model there is no risk adjustment for contra-lateral breast cancer events for trastuzumab treated patients.

The outcome defined by the manufacturer's submission is appropriate for modelling purposes. As explained above, the ERG has no direct way of checking the hazard ratio the manufacturer derives for this outcome. In terms of the published evidence, it should be closest in size to the HERA trial's disease-free survival (DFS). That outcome also includes contralateral breast cancer, second nonbreast malignant disease and death without prior recurrence (totalling, for trastuzumab, 15/127 DFS events; and, for observation, 16/220 DFS events).¹ Along with the timing of events, this could be enough to explain the (CIC data removed) relative difference between the hazard for disease free survival as reported in the HERA publication (HR=0.54) and the hazard for recurrence-free survival in the manufacturer's model input (CIC data removed)

DDFS

The term DDFS refers to "Distant Disease Free Survival" and within the economic model is a variable used to adjust baseline transition probabilities for metastases of the primary cancer in the trastuzumab treated arm. The DDFS variable is essentially the hazard ratio for developing metastases (also called distant recurrence or distant event).

The outcome defined by the manufacturer's submission is appropriate for modelling purposes. The ERG is unclear as to how it differs in its definition from the outcome, 'time to distant recurrence' in the HERA study paper.¹ The HR for this outcome was reported as 0.49 (95% CI 0.38,0.63). The outcome 'DDFS' equals (CIC data removed) in the model. It is unclear from where this discrepancy has arisen.

RFM

The term RFM refers to “Reduction for Metastases”. Within the economic model this is the variable used to adjust the baseline transition probability to metastases for those trastuzumab patients in the recurrence health state. The variable is multiplied to the probability of metastases for patients in the recurrence health state that have received trastuzumab as adjuvant treatment. When analyzing the trial data we found that the risk of developing metastases from the recurrence health state was substantially reduced for trastuzumab treated patients (CIC data removed). Few events and small sample size contributed to this non significance.

Roche have provided sufficient clarification in this area. We have no way of validating this outcome without access to individual patient data. The one year analyses from the HERA trial suggest that this is a rare event as most of the recurrences are metastatic at diagnosis. The extent to which this can be generalised to future years will remain unclear due to the early termination of the study.

A3 Please define what counts as an event (e.g. first or subsequent? Local, regional, distant recurrence, contralateral?) for each clinical outcome used in the model, including whether death without breast cancer recurrence, or the diagnosis of a non-related cancer counts as an event. What matters is how you defined it when you extracted the figures from the trial database: we need a clear and exhaustive definition. It would not suffice to just refer to the publication as it varies in its terminology and is at variance with the ASCO presentation.

The hazard ratios in the clinical trials were calculated for composite end points that included death, second primary cancers etc. Also multiple events were included. This was not appropriate given the required structure of economic model. Therefore a number of re-codings and censorings took place. These are described below.

The ITT population consisted of 1693 patients in the observation arm and 1693 patients in the trastuzumab arm. The economic analysis used the registration data as prepared by Roche Biostatistics in Basel, Switzerland.

Trastuzumab (HERA) Event Censoring Algorithm of Health States

The trastuzumab economic model consists of 6 health states with all patients starting in the Disease Free State (DFS). The six health states defined in the pharmacoeconomic model are:

Disease Free - all patients start in this health state,
Overall Survival (death),
Distant recurrence,
Cardiac event leading to discontinuation of treatment,
Contra-lateral recurrence
Local-Regional recurrence.

Patients transitioning to any one of the remaining 5 health states result in an event being recorded for the patient. The health states are hierarchical in that when more than one event occurs on the same date then the event will be assigned to the event with the highest severity. The order of severity for health states is:

Overall Survival (death)
Distant recurrence
Cardiac Event leading to discontinuation of treatment
Contra-lateral recurrence
Local-regional recurrence

Once a patient has transitioned to one of the five health states then the time to the event is calculated from start of treatment to the date of the event and the censor is set to zero (0). The remaining health states (overall survival excluded) will have the same event time with the censor variable set to one

(1) indicating that for this health state the patient is censored for this health state. Death is the only health state that patients can transition into having first transitioned to one of the other recurrent health states.

Roche have provided sufficient clarification in this area.

A4 Please could you provide the details of the methodology used to calculate the trastuzumab risk reduction ratios and confidence intervals used within the model. The table at the top of the model worksheet: "Adj and Risk Redu Factor" is labeled "risk reduction", therefore our assumption would be that these figures would equal one minus the hazard ratio given in the SABCS abstract. However, if this is the case, the figures do not appear to add up (it may be that you took data directly from the HERA database).

As previously described in question 1 above, the trastuzumab risk reductions are calculated using a standard treatment only adjusted Cox model. The calculations were performed on the HERA registration database but only data on time to first event was used, with first event being specified according to the censoring algorithm explained in question A3 above. Consequently one cannot reproduce the hazard ratios by utilising those figures reported in the main publication.

Hazard ratios were calculated for local recurrence, contra-lateral breast cancer, distant recurrence, and distant recurrence after a local recurrence or contra-lateral breast cancer event.

The Hazard ratio is derived from the fitted Cox's proportional hazards model. Under the Cox proportional hazards model, the hazard function for an individual i can be written as:

$$h_i(t) = \exp^{\beta x_i} h_0(t)$$

This equation is the hazard function (see: Altman DG, *Practical Statistics for Medical Research*, p388;⁵⁶ Collett D, *Modelling Survival Data in Medical Research*, p55, Figure 3.2⁵⁷).

If the treatment is the only covariate, then $h_0(t)$ is the baseline hazard function (hazard function for a patient taking the standard treatment), x is an indicator variable of the treatment administered to the individual i (e.g. 1 for the new treatment, 0 for the standard treatment), and β is a parameter indicative of the treatment effect and is estimated by fitting the Cox model.

The hazard ratio is a measure of the relative hazard (new treatment vs. standard treatment):

$$HR = \exp^{\beta * 1} h_0(t) / \exp^{\beta * 0} h_0(t)$$

$$HR = \exp^{\beta}$$

This equation is the hazard ratio (see: Altman DG, *Practical Statistics for Medical Research*, p388;⁵⁶ Collett D, *Modelling Survival Data in Medical Research*, p56, Figure 3.3⁵⁷).

When fitting the Cox proportional hazards model a parameter estimate for β (b) is derived as well as the associated estimated standard error for b: $se(b)$. These standard errors can be used to obtain an approximate confidence interval for the unknown parameter β . The limits of the confidence interval are the following:

$$b \pm z_{\alpha/2} se(b)$$

where $z_{\alpha/2}$ is the upper $\alpha/2$ - point of the standard normal distribution. A confidence interval for the hazard ratio can be found by exponentiating the confidence limits of the interval for β .

This equation is the confidence interval for the hazard ratio (see: Collett D, *Modelling Survival Data in Medical Research*, p67, Section 3.4⁵⁷).

All of the above methodology which Roche has applied is appropriate.

A5 Please clarify whether the hazard ratios given for disease free survival relate to the actual median one year follow-up, or are estimated for some projected two year follow-up. If possible please provide absolute event numbers and an equation showing how you got to the hazard ratios from the absolute event numbers so that we can determine exactly what these figures represent (relative risks reductions, relative hazard reductions etc).

Hazard ratios for DFS relate to the actual median one year follow-up. The equation for calculating the hazard ratio is described in question A4 above. In relation to providing the raw data, please see question A6 below.

Roche have provided sufficient clarification in this area.

A6 Please provide the absolute event numbers on which these transition probabilities are based and the details of the methodology used to obtain the probabilities. The source of the transition probabilities, given in the economic model sheet "general input" in the table Epidemiological Input is given as the HERA clinical database. It would be useful to have the raw data on which these probabilities were estimated in order to allow us to understand the methodology used & check the calculations.

The methodology for calculating the transition probabilities was based upon the standard formula:

$$Tp = S(t-1) - S(t)$$

The above formula is incorrect. The standard formula here should be $Tp = \{S(t-1) - S(t)\} / \{S(t-1)\}$. However, since the above formula defined by Roche has only been applied where $t = 1$ (because all probabilities are based on the first year's data), the transition probabilities applied within the model are reasonable.

We note your request to provide raw data however Roche does not consider it necessary or appropriate to provide this for the following reasons:

- 1) Any uncertainty relating to the value of these relative risk reductions have been illustrated via comprehensive sensitivity analysis.

The ERG disagrees with the assertion that uncertainty relating to the relative hazard reductions has been adequately dealt with and, as detailed below, the sensitivity analysis is deficient. There is, inevitably, enormous uncertainty when a one year median follow-up analysis of surrogate outcome measures is used to extrapolate to the life-time survival experience.

- 2) The figures, whilst not measuring the exact same events, are consistent compared to the treatment effects (relative risk reductions) observed within the published data.

This is true for the outcomes which have close analogues in the published data as, for example, RFS and DDFS (see responses to Question A2). It is not the case for RFM, although the ERG accepts this is comparatively rare.

- 3) Whilst being a Roche sponsored study, the HERA clinical trial was actually performed by BIG (Breast International Group) and consequently the raw patient level data is not owned or controlled by

Roche. Roche purchased a copy of the data for the purposes of regulatory analysis only.

As detailed in our response to Q1 above, it is impossible for the ERG to validate Roche's analysis without access to the individual patient data. This is not the first time that third party ownership of individual patient data has prevented adequate independent validation of survival and safety outcomes (the recent aromatase inhibitors technology appraisal is another example). It is in the public interest that this issue is addressed in the future. The private ownership of results has deleterious consequences for the public interest in terms of data analysis and interpretation^{58,59}, secrecy⁶⁰ and publication bias⁶¹.

Section B: Economic Sensitivity Analyses

Additional Model Refinement

As part of investigating your questions, making changes to the model and performing the requested sensitivity analysis below, Roche has identified a structural error in the model that was originally submitted to NICE. The error relates to the application of the cost of trastuzumab within the metastatic health state for the no treatment arm of the evaluation. The model only assumed those patients entering the metastatic health state from the recurrence health state were administered trastuzumab. Consequently the total cost within the no treatment arm was underestimated as those patients entering the metastatic health state from the disease-free and cardiac event health states were not given trastuzumab.

The outcome (QALY) element of the no adjuvant trastuzumab arm remains unchanged as it was originally assumed that all of these patients within the metastatic health states received the clinical benefit of trastuzumab. The subsequent impact upon the cost per QALY for trastuzumab was for it to be reduced from the original base case result of £5,700 to £1,900. All subsequent results reported below are now based upon this corrected model version.

The additional model refinement has drawn the ERG's attention to the assumptions surrounding the use of trastuzumab in the Roche model. Roche have assumed:

- (a) All patients in the comparator arm will receive trastuzumab in the metastatic setting.
- (b) Patients who have previously received trastuzumab for primary breast cancer will not receive the drug in the metastatic setting.

With regards to (a), there are likely to be a percentage of patients who will not receive trastuzumab in the metastatic setting regardless of previous chemotherapy regimens. However, the ERG does not anticipate that this will affect the model results significantly.

With regards to (b), without current NICE guidelines suggesting otherwise, the ERG's clinical experts suggest that women who have already been administered trastuzumab for primary breast cancer are likely to receive the drug during metastatic recurrence assuming that they have not been administered trastuzumab for at least around one to two years. Since the model suggests that the average time spent in the recurrence health state is 3.5 years, it is likely that a large percentage of women will receive a second dose if required. In addition, the current NICE guidelines surrounding metastatic breast cancer would suggest that trastuzumab should be administered if appropriate.

Unfortunately, there is currently no evidence to suggest whether trastuzumab is likely to be equally effective in these cases or if resistance may become a problem. However, the ERG feels it is appropriate to assume that *all* patients will receive trastuzumab in the metastatic setting (see section 7.2.1).

B1 A number of extra sensitivity analyses would be helpful for the Committee in their deliberations. Please provide a sensitivity analysis for each of the following individually, and in a combined analysis

together. A more detailed explanation and rationale for these analyses is provided at the end of this letter (Appendix A), but basically we would like to identify the kinds of scenarios that may have a significant effect on the final ICER. These would be used as a basis for debate about which is considered the most realistic.

i) Increased drug cost including cost of administration to £25,000 and £30,000.

Drug costs

From further analysis of the HERA trial data the average number of vials utilised was calculated as 3.33, this figure includes the loading dose and assumes that wastage occurs (i.e. actual vial usage is rounded up prior to calculating the mean). In addition to the assumed average dose which affects the total drug cost, the number of doses was also queried in Appendix A. Further examination confirmed that within a 12 month period, 1 loading dose followed by 17 maintenance doses can indeed be administered. Consequently the revised total drug cost applied within the model for trastuzumab is summarised below:

Table 13: Revised adjuvant trastuzumab drug costs

Assumption	Value
Loading dose	1
Maintenance dose	17
Average number of vials required	3.33
Cost per dose	£1,356.64
Total cost of trastuzumab in adj. setting for 18 doses (1 year)	£24,420

For the metastatic setting, the following trastuzumab drug costs were assumed. These costs represent an increase in £17 over the original submission due to validating the rounding of costs in this calculation.

Table 14: Revised metastatic trastuzumab drug costs

Assumption	Value
Patient weight	70kg
Loading dose	1
Maintenance dose	41
Loading dose (4mg/kg)	280 mg – 2 vials
Maintenance dose (2mg/kg)	140 mg – 1 vial
Cost per Loading dose	£814.80
Cost for 41 maintenance doses	£16,703.40
Total cost of trastuzumab in metastatic setting for 42 doses (42 weeks/approx 1 yr)	£17,518

Unit Cost of Drug administration

As noted in the question above, the cost of drug administration was excluded from both the adjuvant and metastatic (year 1) health states in the original model submitted to NICE. Roche has subsequently added additional model inputs to allow drug administration within the relevant settings to be added if required. The assumed unit cost of a hospital visit for drug administration was calculated as follows:

Table 15: Revised trastuzumab administration costs

Cost	Value	Reference
Cost of complex infusion	£38.00 (2004 prices) £41 (2006 prices)	HTA Review of Guidance No. 33, Addendum on economic evaluation. This cost includes the pharmacist's time for checking and the technician's time for dispensing
Cost of administration	£109 (2004 prices) £117 (2006 prices, used GDP deflators HM Treasury)	HTA review guidance No 33. Taken from PSSRU data. This cost was assumed to include nursing time for the administration of chemotherapy
Total cost for 1 admin	£158	=£117+£41

Total drug administration costs in adjuvant setting

Assuming 18 hospital visits for the administration of trastuzumab within the adjuvant setting generates a total drug administration cost of £2,844 per patient (18*£158). This cost has been added to the model in cell J43 of the "input cost" worksheet.

Total drug administration costs in metastatic setting

To account for drug administration within the metastatic health state, all patients are assumed to receive 6 cycles of conventional chemotherapy within the "Cost of 1st year in the Metastatic State" health state. Consequently a cost of £948 (6*158) has been added to the cost of this health state. The drug costs of conventional metastatic chemotherapy were already included within the metastatic health state costs. For subsequent years in the metastatic setting, 4 cycles per annum are assumed on average; therefore the health

state cost “Cost of one year in the Metastatic State after 1st year” has been increased by £632.

The model currently assumes trastuzumab is only administered in the metastatic setting within the “no adjuvant trastuzumab” arm of the model, assuming 42 weekly visits to administer trastuzumab plus any concomitant chemotherapy. To account for the marginal increase in drug administration cost for those patients receiving trastuzumab in the metastatic setting an additional £5,688 is added to account for the additional 36 hospital visits compared to the 6 visits for conventional chemotherapy. This addition is made within the cell marked “Cost of Trastuzumab Adm for one year of metastatic treatment” in the “Input cost” worksheet of the updated economic model.

Total Drug and Drug administration Costs

Table 16: Revised drug and drug administration costs

	Adjuvant Setting		
	Drug costs	Drug admin	Total
Trastuzumab	£24,420	£2,844	£27,264
No Trastuzumab	N/A	N/A	

To enable a comparison with the costs suggested in question Bi) above; the table above combines the total drug and administration costs assumed within the updated model. After accounting for changes in drug cost and the inclusion of relevant drug administration costs, the figure of £27,264 for trastuzumab falls within the range suggested (£25,000 - £30,000).

The ERG is satisfied with Roche’s new drug cost.
--

Potential Inconsistencies in the model calculations

Appendix B lists possible calculation errors within the model identified by the ERG. Appendix B has been modified to include comments from Roche in

relation to each item. Agreed cost changes listed in appendix B have subsequently been incorporated into the revised model.

The ERG is satisfied with the majority of the comments Roche have made regarding the potential inconsistencies. However, there is one potential inconsistency which we still feel is unresolved:

The monthly cost of endocrine therapy/ chemotherapy for metastatic active treatment was calculated by working out the proportional cost of each which seems reasonable. However, this number was then divided by the total proportion because it was greater than 100% - there is no logical explanation for doing this. However, a sensitivity analysis of this cost suggests that it makes little difference to the ICER.

Revised base case cost per QALY estimate provided in letter from Roche dated 28th March 2006

In summary the following amendments have been made to the economic model and described in greater detail above:

- 1) Updated drug cost for trastuzumab to reflect the mean dose from the HERA trial and extra maintenance dose per year.
- 2) Inclusion of drug administration costs within both the adjuvant and metastatic setting.
- 3) Corrections of cost calculations outlined in table 2 of HTA questions entitled “potential Inconsistencies in the model calculations”.
- 4) Addition of staff costs to FISH test (see question D2 below).
- 5) Structural correction affecting drug costs for those patients receiving trastuzumab within the metastatic setting.

*Consequently the revised base-case incremental cost per QALY for trastuzumab is now **£2,387**.*

The ERG accepts the revisions to the model but considers that the base-case incremental cost per QALY for trastuzumab should be higher than this (see section 7.2.1).

ii) Improving / reducing the survival estimates in the comparator arm

Table 17: Roche's sensitivity analysis around the transition probability from DFS to metastatic

Variable	Value	ICER
Transition probability from DFS to metastatic	<u>(CIC data removed)</u>	£4,756
Transition probability from DFS to metastatic	<u>(CIC data removed)</u>	£1,632

The above table illustrates the change in the incremental cost effectiveness ratio (ICER) when the survival estimates of the comparator arm are varied. The base case value of progressing from DFS to the metastatic health state is approximately **(CIC data removed)**. Consequently the above one-way sensitivity analysis illustrates the scenario where the base case value is adjusted by a factor of approximately +/- 50%. The analysis illustrates the intuitive findings whereby a smaller baseline risk of progression results in a higher cost per QALY for trastuzumab. This is explained by the fact that if applying an identical relative risk reduction to a smaller baseline transition probability, this will result in a smaller absolute clinical benefit for trastuzumab compared to no treatment.

Whilst it is desirable to validate the estimated survival of the comparator arm of the model against published evidence of survival, this is extremely difficult. To achieve this method of validation one would require an estimate of the mean survival for HER2 positive women from the point of early diagnosis/adjuvant setting to death. Firstly this would require an impractically

long prospective study due to the long life-expectancy of these patients. Secondly HER2 testing is only a relatively recent practice and therefore it is highly unlikely any such long term evidence of HER2 positive survival from the adjuvant setting is available. The articles referred to by the HTA in relation to this issue do not appear provide evidence to perform such a validation.

However Figure 1 of the Stal et al (2000) paper referred to below illustrates graphically that approximately 60% of HER2 patients are alive after 10 years. The Roche economic model estimates that approximately 63% of HER2 positive patients are alive after 10 years in the comparator arm. Although not a highly robust method of validation, this provides a reassuring comparison of baseline survival estimated by the model.

The above sensitivity analysis was considered to be inadequate. Roche appear to have chosen one event (DFS to metastatic recurrence) and carried out a one-way sensitivity on that transition probability whilst ignoring all others. This does not give a sense of the effects of a better or worse comparator arm upon the ICER as was required.

Therefore, the ERG has carried out their own sensitivity analysis around the comparator arm (see section 7.2.2.3).

iii) Sensitivity analysis around extrapolation of benefits of trastuzumab beyond the trial period, by setting the relative risk of trastuzumab equal to 1 after (a) 10 years (b) 5 years.

Table 18: Roche's sensitivity analysis around extrapolation of benefits of trastuzumab

Variable	Value ¹	ICER
(a) Trastuzumab risk reduction From 11 - end	RFS (Mean) – 1 DDFS (Mean) – 1 RFM (Mean) - 1	£2,792
(b) Trastuzumab risk reduction From 5 – 10 yrs	RFS (Mean) – 1 DDFS (Mean) – 1 RFM (Mean) – 1	£4,461
Trastuzumab risk reduction From 11 – end	RFS (Mean) – 1 DDFS (Mean) – 1 RFM (Mean) - 1	

The above analysis illustrates the cost per QALYs for trastuzumab where no relative risk reduction is assumed for trastuzumab beyond year 11 and beyond year 5 within the model. To simulate this effect the variables RFS, DDFS and RFM are all set to 1 within the model. In both scenarios the cost per QALY increases very little. This may be partly explained by the influence of discounting upon the QALYs; any advantages in the treatment affect of trastuzumab assumed in later years actually receives very little weight in the base case.

Roche have carried out the analysis which the ERG requested. However, since the ERG have suggested that trastuzumab is likely to be given to the majority of patients in the metastatic setting which has a profound effect upon the ICER, the sensitivity analysis above has been repeated with the inclusion of the new assumption (see section 7.2.2.3).

¹ RFS is Recurrence Free Survival, DDFS is Distant Disease Free Survival and RFM is Reduction For Metastases. For more details see above, Section A2.

iv) Widening the confidence intervals around RFS, DDFS and RFM curves over time to reflect uncertainty in these measures.

In order to demonstrate the impact of this assumption upon the final cost per QALY it is necessary to re-run the probabilistic sensitivity analysis. Consequently, illustrated below are 2 cost effectiveness acceptability curves (CEACs), summarising the PSA results for the base-case cost per QALY and the amendments to the confidence intervals requested in question iv) above.

The assumed confidence intervals around the base-case relative risk reductions are already notably wide. However at the request of the HTA these have been widened yet further in years 5-10 and 11+ and are illustrated below:

Table 19: Roche's sensitivity analysis of the confidence intervals used in the extrapolation of the trastuzumab arm

	RFS		DDFS		RFM	
	Lower	Upper	Lower	Upper	Lower	Upper
1-5	<u><i>(CIC data removed)</i></u>					
5-10	<u><i>(CIC data removed)</i></u>					
11+	<u><i>(CIC data removed)</i></u>					

Probabilistic Sensitivity Analysis

Figure 3: Cost effectiveness acceptability curve, base case

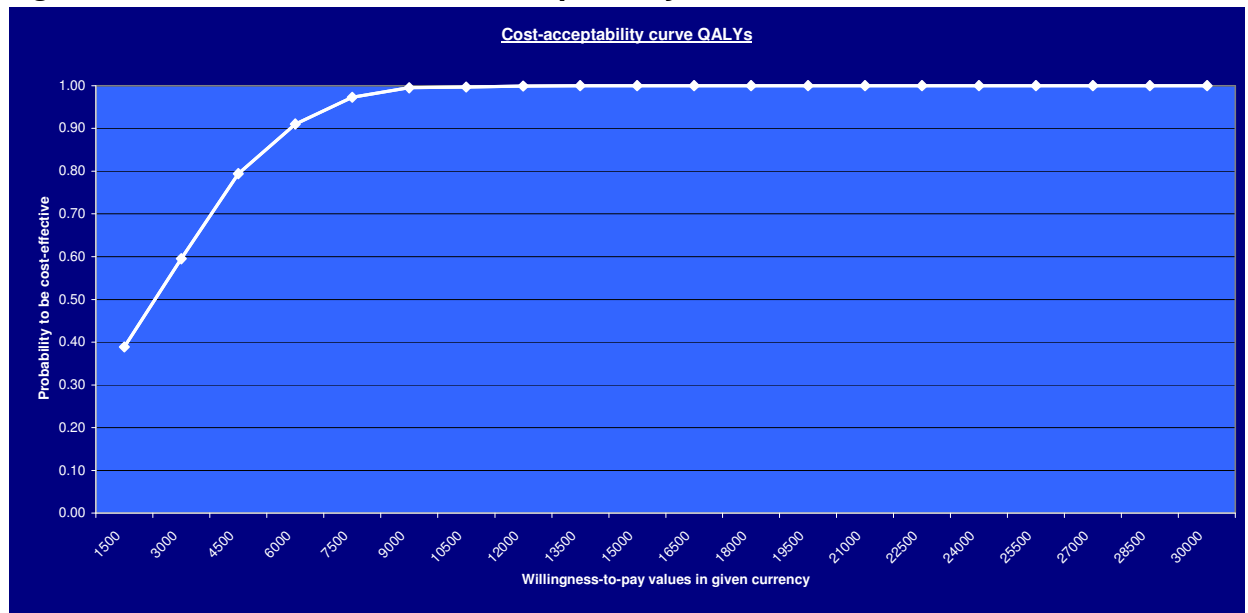


Figure 3 illustrates the probability that trastuzumab (based upon base case cost per QALY £2,387) is considered cost effective across a range of potential cost effectiveness thresholds. This CEAC illustrates that at a willingness to pay of £20,000 the addition of trastuzumab is cost effective in all cases. Only at a threshold value of less than £2,000 does the probability that trastuzumab can be cost effective fall below 50%.

Figure 4: Cost effectiveness acceptability curve, widened confidence intervals

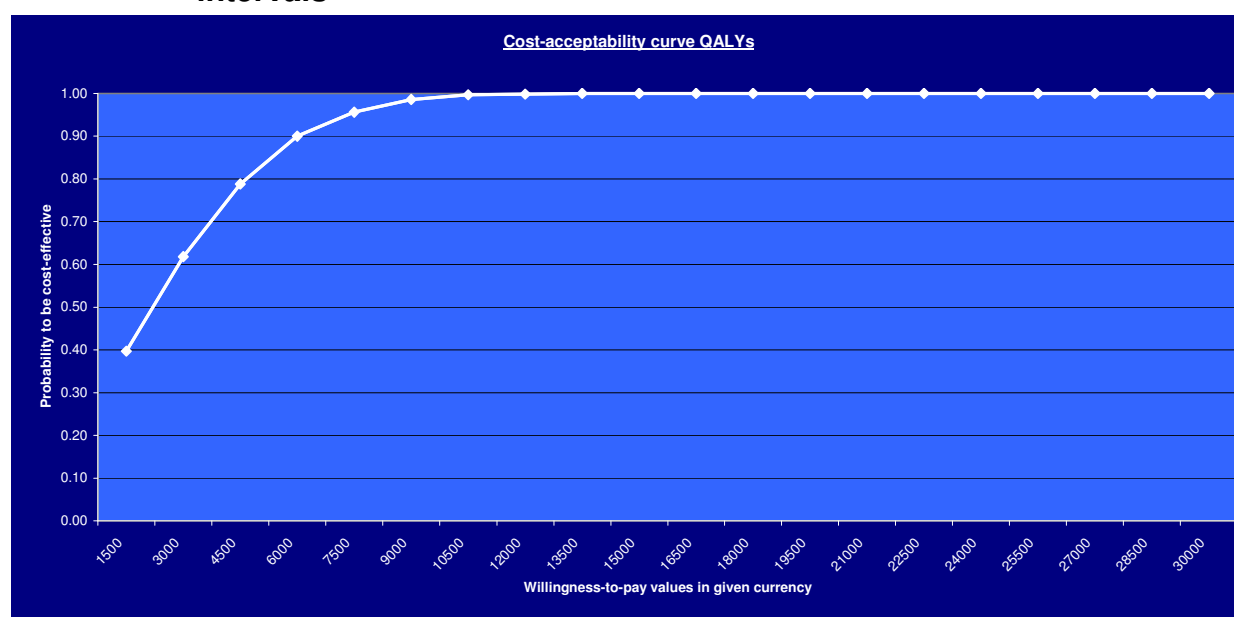


Figure 4 illustrates the probability that trastuzumab is considered cost effective across a range of potential cost effectiveness thresholds, assuming an increasing uncertainty around the relative risk reduction values. The figure illustrates very little change despite this assumed increase in uncertainty. The CEAC illustrates that at a willingness to pay of £20,000 the addition of trastuzumab is cost effective in all cases.

Section C: Effectiveness Data

C1 Please provide us with a summary of FINHER that corresponds with the section 2.3, 2. and 2.5 of the submission template. The results of the FINHER study are not drawn on extensively in the submission because it compares “two chemotherapy regimens neither of which would be much used in the UK.” However you indicate in your submission that the effect of trastuzumab appears consistent regardless of the comparator.

We note that the trastuzumab dosing schedule in FINHER is not the same as in the HERA study, and therefore these other studies may be seen as less

relevant to the Committee's discussion. However without a final licensed indication and dosing regimen we feel that the results could be important.

As we explained in our original submission, the FINHER study is not a Roche study and our knowledge of it is limited to what is in the public domain. At the time of submission, this was limited to the two abstracts of meeting presentations by Joensuu et al. copies of which were included in our recent submission. These have since been supplemented by a peer-reviewed publication (Joensuu et al. 2006⁴). We would refer you to these publications. However, we would draw attention to the following points to be considered when reviewing this study:

The study was primarily an adjuvant chemotherapy study randomising patients to 3 cycles of docetaxel or vinorelbine followed by 3 cycles of FEC. Neither vinorelbine nor docetaxel are much used as adjuvant treatments in the UK and vinorelbine is not licensed in this setting. Whilst HERA shows that trastuzumab can beneficially be added to a variety of chemotherapy regimens (and there is no reason to believe that those used in FinHER are exceptions), toxicity data (vital to this appraisal and likely to be influenced by choice of cytotoxic regimen) derived from a study using only atypical chemotherapy regimens needs to be treated with caution.

Trastuzumab was administered weekly at the same time as docetaxel/vinorelbine and before the anthracycline portion of chemotherapy. In HERA, trastuzumab was administered three-weekly once all chemotherapy had finished. Again this might be expected to influence the toxicity of treatment.

The sub-study within FinHER looking at the impact of trastuzumab on women with HER2 positive breast cancer included only 232 patients of whom just 115 received trastuzumab and, by the time of analysis only 39 DFS events had occurred. Thus, encouraging though the results are in terms of trend, this data is not a secure basis for decision making.

Most crucially, although the final licensed indication and dosage regimen for adjuvant trastuzumab is not yet confirmed, the Roche regulatory dossier is

based on the HERA dosing schedule and data. The EMEA have the option to accept or reject our data and our proposed indication and posology, but it is not within the realms of possibility that they will impose a dosing strategy completely different from that used in our regulatory study.

Roche have not met the ERG's request to include a summary of FinHer that corresponds with sections 2.3, 2.4 and 2.5 of the submission template. The ERG recognises that the FinHer trial evidence must be approached with caution (See this report, Sections 4.1.3.2 and 7.1.5), however a summary of the evidence would have been considered useful as part of the report.

C2 If you do not feel this is appropriate then you may want to provide further rationale for not drawing on this data.

See response to C1

C3 Please provide a table with baseline characteristics for the trials other than HERA. Section 2.3.2 only provides baseline characteristics for HERA. Although it is appreciated that only limited information may be available for BCIRG006, NEJM articles are available for B-31, N9831 and FINHER.

These are not Roche studies and we have no further information on the baseline characteristics of entrants to BCIRG 006, NSABP B-31, NCCTG N9831 and FinHER beyond that available in the public domain. In the case of the last 3 studies detailed information is provided in tabulated form in the NEJM articles of which you already have copies and we would refer you to these.

The ERG suggests that Roche should have included this information within the report to provide a more complete overview of the effects of trastuzumab.

C4 Please provide a meta-analysis of the results for HERA and the combined results from B-31 & N9831, BCIRG 006 and FINHER. The submission states that there is a consistency of effect across the trastuzumab studies. If this is correct, then there doesn't seem to be a reason why such a meta-analysis cannot be executed in Review Manager or other such software. This analysis would help validate the claims of effectiveness regardless of comparator, and provide an easy way for the NICE committee to interpret and understand the effectiveness data.

As we explained in answer to Q58 in our original submission, we do not see the value on this occasion of undertaking any additional meta-analysis when several large studies all reach essentially the same conclusion. Amalgamating the data from these studies would not clarify an area of uncertainty but might submerge important differences between studies that result from differences in their designs.

A meta-analysis of the data would provide a clear comparison with the other trials and would have helped to validate the data from the HERA trial to provide a stronger evidence base. The ERG has carried out a meta-analysis of the outlined trials, section 7.1.

Section D: Additional Discussion and Rationale

D1 Based on the figures presented in the “validation” sheet of the submitted model it can be seen that the patients in the comparator arm in the USA trials have a better prognosis than the patients in the comparator arm from the HERA study (84.8% in the joint US analysis compared with 77.4% from the HERA trial). A more detailed discussion of this issue and why these differences exist would be most useful to the Committee in its discussions.

Cross-trial comparisons are fraught with difficulty and there are a large number of subtle differences between the patients recruited into HERA and

US studies that could explain the small differences in the performance of the control arms between the studies. However, in this case the most likely explanation is that the standard adjuvant therapy onto which the trial intervention has been imposed is different between HERA and the US studies.

All patients in the joint analysis of B-31 and N9831 received an anthracycline and a taxane as part of their adjuvant chemotherapy compared with only 25.6% of control patients in HERA. As explained in our original submission, although taxanes are not widely used in the UK as part of adjuvant chemotherapy there are data from at least 3 randomised trials (reviewed in Smith, 2005 as previously supplied) that they increase DFS by 4%-7% in absolute terms. This alone could account for the difference in outcomes in the control arms of HERA and the US studies.

In addition, more control patients in the combined US studies (19.4%) received aromatase inhibitors as part of their hormonal treatment for hormone receptor positive tumours than did those in HERA (10%). Since aromatase inhibitors have been shown to improve DFS in absolute terms by 3-5% when substituted for or added to tamoxifen (again, reviewed by Smith, 2005), the increased use of aromatase inhibitors by North American patients is likely to have contributed towards the improved outcome amongst control patients in US studies.

The increased use of taxanes and aromatase inhibitors together is quite sufficient to explain the better outcome amongst control patients in the US. However, it must also be noted that the reduction in the risk of relapse is similar in the US and HERA studies i.e. there is no reason to think that the risk reduction mediated by trastuzumab will diminish if the UK moves towards a more American pattern of adjuvant therapy.

Whilst the above explanation is reasonable, the ERG suggests that a sensitivity analysis of the comparator arm is essential, particularly given that the baseline Roche have assumed may change if NICE recommend the use of taxanes for high risk patients with early breast cancer in the near future (see section B).

D2 HER 2 testing: Initial examination of the model by SchARR has suggested that the cost of HER2+ testing may be underestimated. Although at this stage it is not considered that the cost of these tests will make a major difference to the results of the modelling work, issues around the sensitivity and specificity of the tests and the impact of this on the potential for inappropriate treatment of patients have not be discussed. Can you add a discussion of your views on the issues relating to the accuracy of HER2+ testing and the likely impact on treatment costs for breast cancer patients.

Cost of testing

We assume the underestimate of the HER2 testing costs described above relates to the fact that no staff costs were included as part of the cost of FISH testing. Consequently this cost has been increased by £35 to capture staff costs (medical solutions plc, 2004). We assumed that all early breast cancer would be tested by IHC, with FISH testing used subsequently for the 10% of patients whose tumours give a 2+ score on IHC. This rate of FISH testing was confirmed as appropriate with organisers of the National Quality Control Scheme for HER2 Testing (NEQAS).

Accuracy of testing

The testing schedule used in HERA was essentially the same as that used in UK clinical practice: IHC testing with FISH testing used on borderline IHC 2+ samples. Therefore, if testing is rigorously conducted there should be no concern that false positive or negative results are any more or less likely in clinical practice than in HERA.

In fact, UK pathologists were quick to recognise the importance of a rigorous approach to the standardisation and quality control of HER2 testing by both IHC and FISH techniques and national guidelines exist to ensure the quality of HER2 testing (Ellis et al. references 14 and 15 in our original submission). The combination of these guidelines, the existence of a nationally organised quality assurance scheme and the availability of FISH testing as an additional discriminator where IHC testing (generally used first-line) is equivocal mean

that there is no reason to doubt that high levels of accuracy are achieved, as verified by NEQAS.

Additionally, following the announcement made by the Secretary of State for Health that all women with breast cancer should be HER2 tested at diagnosis, Roche has undertaken a project in conjunction with the National Cancer Director to ensure all Cancer Networks have the necessary infrastructure in place to allow this to occur reliably. The project, described on page 105 of our original submission, is designed to ensure that the resources are in place to ensure that the following aspects of HER2 testing can be carried out efficiently and to a high standard:

Pathologist resources

Her-2 testing kits and consumables

Block preparation

Slide preparation

Staining

FISH testing (in 10 – 15% of patients)

Reading of results

Reporting of results.

It should be noted that HER2 status is an important prognostic marker and it is not carried out solely for the purposes of deciding whether trastuzumab should be given. As such, it is probably inappropriate to link all of the costs of this procedure to one intervention, though this is what Roche has done in our economic model.

The ERG is satisfied with Roche's response and revised cost of HER2 testing.
--

7 Additional work undertaken by the ERG

7.1 *Clinical effectiveness*

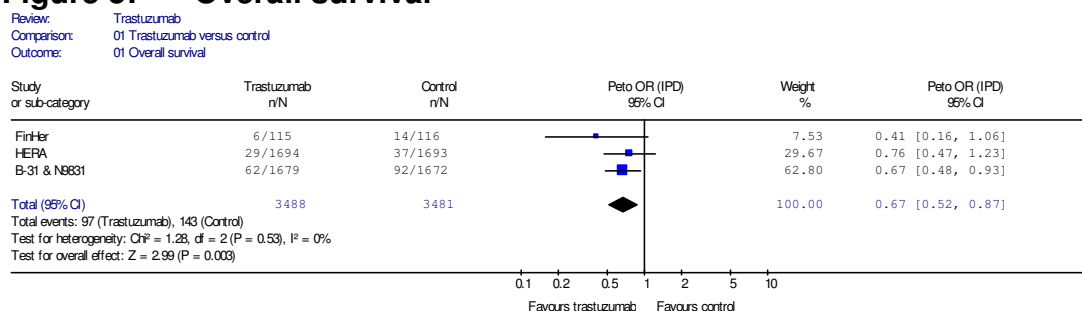
In this section we carry out the following analyses which the manufacturer declined to undertake: (1) a meta-analysis of trials to derive a more precise estimate of treatment effect in terms of overall survival (Section 7.1.1), disease-free survival (Section 7.1.2), distant recurrence (Section 7.1.3) and cardiac toxicity (Section 7.1.4); and, (2) a critical evaluation of the role of the FinHer study in decision-making (Section 7.1.5).

For time-to-event outcomes, summary statistics from the published literature were meta-analysed using the method described by Parmar⁶² with a fixed effects model. Heterogeneity between trial results was tested using the chi² test and the I^2 measurement. The chi² test measures the amount of variation in a set of trials. Small p values ($p < 0.10$) suggest that there is more heterogeneity present than would be expected by chance. I^2 is the proportion of variation that is due to heterogeneity, rather than chance. Large values of I^2 suggest heterogeneity. I^2 values of 25%, 50%, and 75% could be interpreted as representing low, moderate, and high heterogeneity.⁶³

The Absolute Risk Reduction (ARR) and Numbers Needed to Treat for time-to-event outcomes were calculated using methods described by Altman and Andersen.⁶⁴ This method uses the numbers of patients still at risk (alive) at the time corresponding to the estimated probabilities (reported or imputed), or hazard ratios and 95% confidence intervals, to calculate confidence intervals for each statistic.

7.1.1 Overall survival

The necessary summary statistics were available to meta-analyse data from four studies: FinHER,⁴ HERA¹ and the combined B-31/N-9831 analysis.²⁴

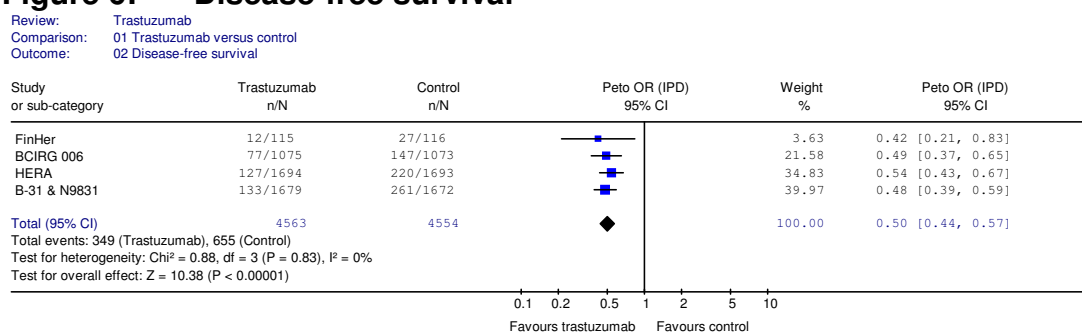
Figure 5: Overall survival

Across four studies (6,969 women) with median follow-ups of between one and four years, the addition of between nine and 52 weeks' trastuzumab to standard care conferred a 33% relative improvement in overall survival (HR 0.67, 95% CI 0.52-0.87, $p=0.003$). There was no statistical heterogeneity between studies.

When analysis was restricted to studies which gave 52 weeks' trastuzumab, as in the manufacturer's proposed marketing authorisation, trastuzumab conferred a 30% relative improvement in overall survival (HR 0.70, 95% CI 0.53-0.92, $p=0.010$). The apparent poor performance of the HERA study compared to the combined B-31/N-9831 analysis is most likely to be a function of short follow-up with, fewer analysed participants and accumulated events contributing information in the former.

7.1.2 Disease-free survival

The necessary summary statistics were available to meta-analyse data from five trials: FinHER,⁴ HERA,¹ BCIRG-006³³ and the combined B-31/N-9831 analysis.²⁴

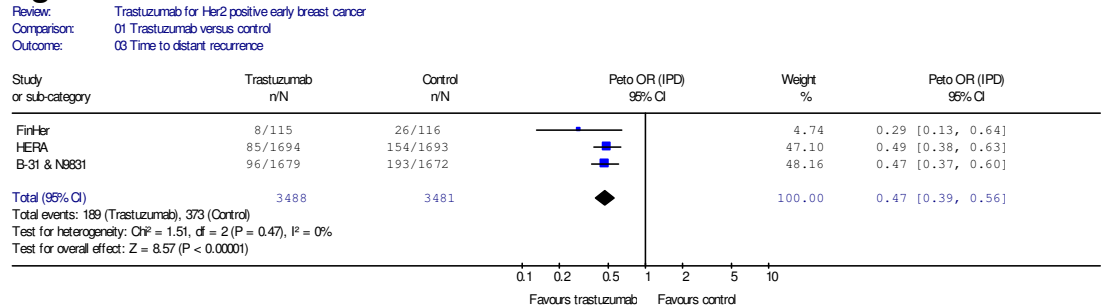
Figure 6: Disease-free survival

Across five studies (9,117 women) with median follow-ups of between one and three years, the addition of between nine and 52 weeks' trastuzumab to standard care conferred a 50% relative improvement in overall survival (HR 0.50, 95% CI 0.44-0.57, $p < 0.00001$). There was no statistical heterogeneity between studies. When analysis was restricted to studies which gave 52 weeks' trastuzumab, as in the manufacturer's proposed marketing authorisation, the summary statistics were unchanged.

7.1.3 Distant recurrence

The necessary summary statistics were available to meta-analyse data from four studies: BCIRG-006³³ HERA¹ and the combined B-31/N-9831 analysis.²⁴

Figure 7: Distant recurrence



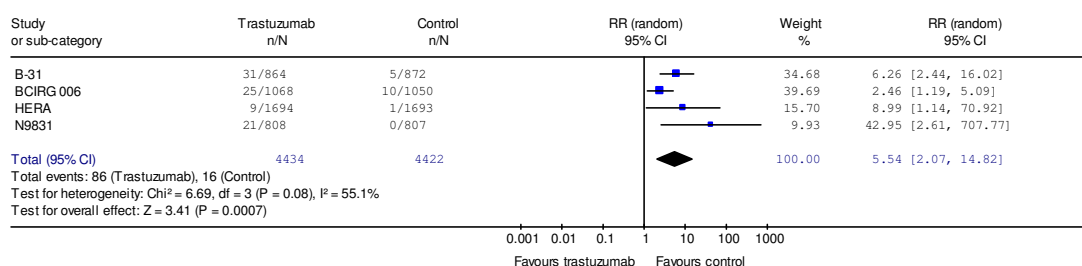
Across four studies with median follow-ups of between one and three years, the addition of between nine and 52 weeks' trastuzumab to standard care conferred a 53% relative improvement in distant recurrence (HR 0.47, 95% CI 0.39-0.56, $p < 0.00001$). There was no statistical heterogeneity. When analysis was restricted to studies which gave 52 weeks' trastuzumab, as in the manufacturer's proposed marketing authorisation, the summary statistics were unchanged (HR 0.48, 95% CI 0.40-0.57, $p < 0.00001$).

7.1.4 Cardiac toxicity

The necessary summary statistics were available to meta-analyse data from four studies: FinHER,⁴ HERA¹ and the combined B-31/N-9831 analysis.²⁴

Figure 8: Cardiac toxicity

Review: Trastuzumab for Her2 positive early breast cancer
Comparison: 01 Trastuzumab versus control
Outcome: 04 Grade 3-4 cardiac event or death from heart failure



Across four studies with median follow-ups of between one and two-and-a-half years, the addition of 52 weeks' trastuzumab to standard care increased the relative risk of a serious, life-threatening or fatal cardiac event almost sevenfold (relative risk [random effects model]=5.54, 95% CI 2.07-14.82, $p=0.0007$). Across all studies, this relative increase in such events of over 500% represents an absolute increase of only 1.6%.

Although treatment effect was statistically significant within each trial, there was moderate statistical heterogeneity between trials ($\chi^2=6.69$, $df=3$, $p=0.08$; $I^2=55.1\%$) indicating potential differences in the underlying interventions (concurrent chemotherapy) and outcome assessments (length of follow-up) which make interpretation problematic (see also, Section 6, Question C4).

7.1.5 Interpreting the results of the FinHer study

The FinHer study has generated considerable interest since its findings were presented at the San Antonio Breast Care Symposium in 2005, with speculation that its shorter treatment schedule is still effective and “may facilitate lower cost, greater patient convenience, and reduced risk of cardiotoxicity.”¹⁰ We have attempted to redress the questionable exclusion of this study (see Section 3.2) from the manufacturer's submission by its inclusion in this report, both in tables and meta-analyses. But, while its results are striking, it is important that they are interpreted with caution. This section discusses what can be understood with any degree of confidence from the results of FinHer in terms of: overall survival (Section 7.1.5.1); disease-free survival (Section 7.1.5.2); and, cardiotoxicity (7.1.5.3).

7.1.5.1 Overall survival

With a median follow-up of three years, the FinHer trialists found that nine weeks' trastuzumab had conferred no significant survival advantage. Even with a longer follow-up than any other study (survival advantages become more apparent over time), this is not surprising: neither the FinHer study's main comparison (docetaxel versus vinorelbine) nor any other included study were powered to detect differences in overall survival (see above, Section 3.4, on the reasons why). If we take at face value the overall survival (death being a relatively rare event in the short term) results from FinHer's small size, we run the risk of a type II (beta) error: the assumption of no survival advantage from trastuzumab when, in reality, one exists.

7.1.5.2 Disease-free survival

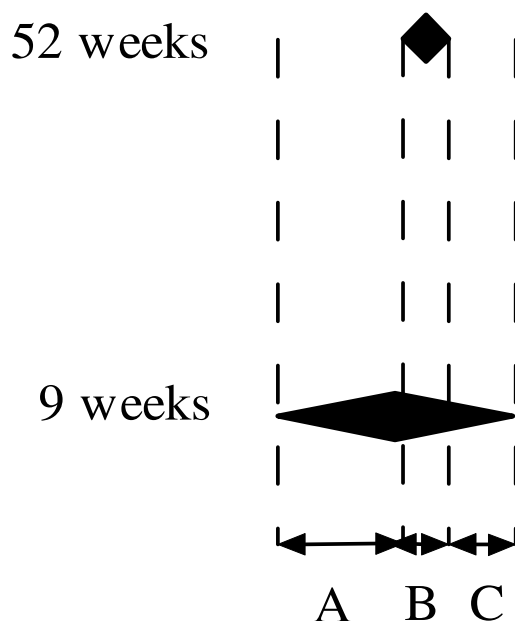
The p value for the disease-free survival treatment effect estimate in the FinHER trial ($p=0.01$) suggests that the null hypothesis ('there is no difference between nine weeks' trastuzumab and observation') is very unlikely to hold. Its proper interpretation is that the observed difference in treatment effects would be found by chance alone once in one hundred times. So, the direction and size of the effect for disease-free survival for the FinHer study (Figure XX) does support the hypothesis that some trastuzumab is better than none. The 95 percent hazard ratio for recurrence or death in the trastuzumab group, as compared with the control group was 0.42 (95% CI 0.21-0.83; $p=0.01$), representing a 58% relative risk reduction (against no trastuzumab) over an average of three years.

However, there is a one in twenty chance that this result is a false positive and that the 'true' effect lies somewhere outside of this range (at either end). Furthermore, as illustrated in Figure 6 above, the wide confidence interval tells us that the estimate of the effect size in FinHer is less precise than those in the other studies. There is a 95 percent chance that the true effect size lies between a 17% and 79% relative risk reduction at this time point. This is critical because, while the central estimate of effect for FinHer ($HR=0.42$) makes its nine week schedule looks more impressive than the 52 week

schedules evaluated elsewhere, the confidence interval includes values which are either side of the confidence interval for 52 week schedules. This is illustrated in Figure 9 below: if the 'true' value is in band 'A', then nine weeks' trastuzumab is more effective than 52 weeks'; if the 'true' value is in band 'B', then the effect is equivalent; but, if the 'true' value is in band 'C', then nine weeks' is inferior to 52 weeks' trastuzumab.²

To emphasize the caution that is needed when interpreting these results, the 99 percent confidence interval for disease-free survival in FinHer is 0.17 to 1.04. This means that, with a greater degree of assurance than typically offered, the range within which the 'true' value lies includes the possibility that 9 weeks' trastuzumab is no better than observation. This is not the case for the pooled estimate of effect from four studies evaluating 52 weeks' trastuzumab (99% CI 0.42, 0.60), nor indeed for any of the individual studies.

Figure 9: 52 weeks' versus 9 weeks' trastuzumab



² Note that the comparatively long follow-up time of FinHer relative to the other studies (three years versus one to two years), also makes interpretation problematic. By analogy with other early breast cancer therapy trials a greater effect size would be expected over time in a truly effective treatment (see, for instance, the Oxford Overview¹³). On the other hand, the very fact that events and censoring occur over time means that confidence intervals also get wider, because fewer people are contributing information to the outcome assessment.⁵⁷ All of this said, most of the uncertainty about the precision of the FinHer effect size must be attributed to the small sample size which the study started out with.

Another method of characterising the robustness of the evidence is to calculate Rosenthal's 'file drawer number'. This is a sensitivity analysis which estimates the number of unpublished neutral trials of equivalent size and variance that would be required to make the result no longer statistically significant.⁶⁵ For the 52 week regimen, 81 hypothetical neutral studies (average event rate: 161 per arm; average trial size $n=2,962$; and, average variance: 72.06) would be needed before our pooled disease-free survival result was overturned. For the FinHer study, only one equivalent-sized neutral study would be required to do the same.

7.1.5.3 Cardiotoxicity

No cardiac events were recorded in the trastuzumab arm of the FinHer trial. While some have speculated that this indicates a real reduction in (or even elimination of) toxicity relative to 52 weeks' treatment, there is again, the risk of a type II (beta) error, with the small sample underpowered to detect significant differences between arms in relatively rare cardiac events (see Section 7.1.5.1).

7.1.5.4 Conclusion

The results of FinHer provide relatively convincing evidence that nine weeks' trastuzumab is better than none at all, and are enough to generate an important hypothesis: that 9 weeks' is equivalent (or non-inferior) to 52 weeks' trastuzumab. But this is *only* a hypothesis, which needs testing, because women are unlikely to value extra convenience at the expense of clinical effect, and a loss of effect size would also have an unknown impact on cost-effectiveness. The worst case scenario is that 9 weeks' treatment prevents fewer breast cancer recurrences than 52 weeks' and is equivalent in toxicity. This outcome seems unlikely, but the confidence intervals allow for this possibility, and decisions must be made accordingly.

7.2 **Cost-effectiveness**

As a result of the communication with Roche, the ERG has developed what they believe to be a reasonable revised base-case. Sensitivity analysis has

also been carried out to ensure that the model results are robust. This analysis is described below.

7.2.1 The ERG's revised base case

7.2.1.1 Provision of trastuzumab in the metastatic setting

In the economic model, all patients in the comparator are assumed to receive trastuzumab in the metastatic setting. Current NICE guidelines recommend that to be eligible for trastuzumab first line anthracycline therapy should be inappropriate or that the patient should have failed two other chemotherapy regimens. It therefore seems unlikely that all these patients would receive trastuzumab and that some may only receive a limited course.

The Roche model assumes that patients receiving trastuzumab for early breast cancer will not receive it in the metastatic setting. Our clinical advisors suggest that in practice trastuzumab may well be offered to these patients if there has been a reasonable interval between the completion of treatment for early cancer and presentation with recurrence. The definition of reasonable is likely to vary between clinicians but may be as short as one or two years.

The ERG have assumed that all patients would receive trastuzumab in the metastatic setting, whether or not they received it for early breast cancer. It is recognised by the ERG that not all patients will be eligible to receive trastuzumab for metastatic cancer therefore this assumption is considered conservative. This increases the ICER from the Roche revised basecase of £2,387 to **£8,365**.

7.2.1.2 Long-term relative risk of recurrence for trastuzumab patients

The benefits for patients on trastuzumab beyond the one year treatment period are not possible to ascertain based on current evidence. Evidence from the joint US trials for trastuzumab, with a median follow-up of two years, suggest that benefits continue for at least two years and possibly three or four years from the start of treatment. The EBCTCG overview¹³ supports the view

that, for some drugs at least, the risk of recurrence is reduced well beyond the treatment period. For instance 6 months of anthracycline-based chemotherapy for patients with early breast cancer provides benefits in terms of a reduction in the risk of recurrence for around five years and the benefits of 5 years treatment of tamoxifen are shown to reduce the risk of recurrence for around 10 years, demonstrating a protective “carry-over” effect for five years beyond the treatment period.

Since there is no evidence of the effects of trastuzumab in early breast cancer after this time, the ERG suggests that a conservative view of there being no further benefit in risk of recurrence after five years should be taken as the base-case. This would result in an ICER of **£4,461**. Applying this assumption, the overall survival also follows a similar pattern to that of the EBCTCG data.¹³

ERG's basecase - combining assumptions from 7.2.1.1 and 7.2.1.2

Assuming that trastuzumab would be provided for all patients in the metastatic setting in addition to there being no further benefit beyond five years, the ICER would become **£18,449**.

Therefore, the ICER for the ERG's base case is **£18,449**. All subsequent sensitivity analyses in this section are based upon this figure.

7.2.2 Sensitivity analysis around the ERG's base case

7.2.2.1 Resistance during provision of trastuzumab for metastases

Because there are often issues with resistance upon providing the same drug twice, a sensitivity analysis has been carried out to consider the possibility of trastuzumab being less effective upon its second administration for those patients who had previously received it in the early breast cancer setting. Reducing the effectiveness of the drug by half during the metastatic health state of the trastuzumab arm would produce an ICER of **£17,905**.

7.2.2.2 Assuming no wastage of trastuzumab

The Roche model assumes that any remaining trastuzumab within the vials for each person will be wasted. However, it may be possible to make use of the remainder of each for the next person. Assuming that each person weighs an average of 70kg, the cost of the total number of vials required per person would be £20,913.20. This reduces the ICER to **£15,828**.

7.2.2.3 Relative risk of recurrence for trastuzumab patients

The ERG's revised base case assumes that there will be no further benefit in terms of risk of recurrence beyond five years. Until this time the relative risk of recurrence is assumed to remain constant. Given that evidence on the relative risk of recurrence for patients on trastuzumab is limited beyond two years a sensitivity analysis to consider a reduction in relative risk of recurrence after two years was carried out by adjusting the relative risk reduction of the trastuzumab arm by **(CIC data removed)** after for years 3 to 5 (See Table 20 below).

Table 20: ERG sensitivity analysis – reduction in the benefits during the first five years following start of trastuzumab therapy

Variable	Assumed Value ³	ICER
Trastuzumab risk reduction from 1 – 2 years	RFS (Mean) <u>(CIC data removed)</u> DDFS (Mean) <u>(CIC data removed)</u> RFM (Mean) <u>(CIC data removed)</u>	£23,566
Trastuzumab risk reduction from 3 - 5 years	RFS (Mean) <u>(CIC data removed)</u> DDFS (Mean) <u>(CIC data removed)</u> RFM (Mean) <u>(CIC data removed)</u>	

The ERG's clinical experts highlight the current uncertainty around whether trastuzumab treatment reduces the number of recurrences or merely delays them. A sensitivity analysis has been carried out to analyse the effect of this upon the ICER (see Table 21 below). It is assumed that in years 5 to 10 the

³ RFS is Recurrence Free Survival, DDFS is Distant Disease Free Survival and RFM is Reduction For Metastases. For more details see Section 6 (A2).

risk of recurrence is higher in the trastuzumab arm than the comparator arm – the relative risk reduction is set at 1.2. Beyond year 10 the risk of recurrence is assumed to be the same in both arms.

Table 21: ERG sensitivity analysis – higher risk of recurrence in years 5 to 10 for patients receiving trastuzumab

Variable	Assumed Value	ICER
Trastuzumab risk reduction from 5 – 10 years	RFS (Mean) – 1.2 DDFS (Mean) – 1.2 RFM (Mean) – 1.2	£23,256
Trastuzumab risk reduction from 11 - end	RFS (Mean) – 1 DDFS (Mean) – 1 RFM (Mean) – 1	

In addition, Roche assumed that the time spent in local or contralateral recurrence would be increased before metastatic recurrence following trastuzumab provision. This was based on a small number of patients from the HERA trial, however because only a small sample reached this stage during the follow-up period, this effect was statistically insignificant. Therefore, the ICER has been recalculated based on an alternative assumption that the time between local or contralateral recurrence and metastatic recurrence is the same for both arms (see Table 22 below).

Table 22: ERG sensitivity analysis - Time between local/contralateral recurrence and metastatic recurrence equivalent for both arms

Variable	Value	ICER
Trastuzumab risk reduction from 1 – 5 years	RFS (Mean) – <u>(CIC data removed)</u> DDFS (Mean) – <u>(CIC data removed)</u> RFM (Mean) – 1	£18,643

7.2.2.4 Sensitivity analysis around overall survival

Table 11, section 5.1.1 suggests that the overall survival in the comparator arm is overestimated, determined by transition probabilities. The majority of transition probabilities are extracted from the HERA trial data, although the probability of transition from metastatic disease to death is taken from a previous Roche paper around the use of trastuzumab for metastatic breast cancer.⁴² This assumption is tested in a sensitivity analysis around the probability of moving from local or contralateral recurrence to metastatic recurrence since these are the two parameters which will reduce the overall survival of both arms (see Table 23 below). The first sensitivity analysis is based upon the transition probabilities used within SchARR's hormonal therapies model, while the second multiplies each of the transition probabilities by (CIC data removed) such that the survival curve produced is in line with the survival indicated by the research papers outlined in Table 11, Section 5.1.1.

In addition, the benefits of trastuzumab in terms of overall survival are not yet known due to the short follow-up data available. Therefore, using the above transition probabilities, the parameter for the relative risk of recurrence from local to metastatic recurrence has been altered from (CIC data removed) to 1 as in Table 22 above. The combined effect of this upon the ICER is shown in brackets below.

Table 23: Decrease in overall survival

Transition probability	Initial value	Values based on SchARR's hormonal therapies model	ICER	Values multiplied by <u>(CIC data removed)</u>	ICER
Local recurrence to metastatic recurrence	<u>(CIC data removed)</u>	0.117	£18,488 (£18,686)	<u>(CIC data removed)</u>	£18,579 (£18,776)
Metastatic to death	<u>(CIC data removed)</u>	0.373		<u>(CIC data removed)</u>	

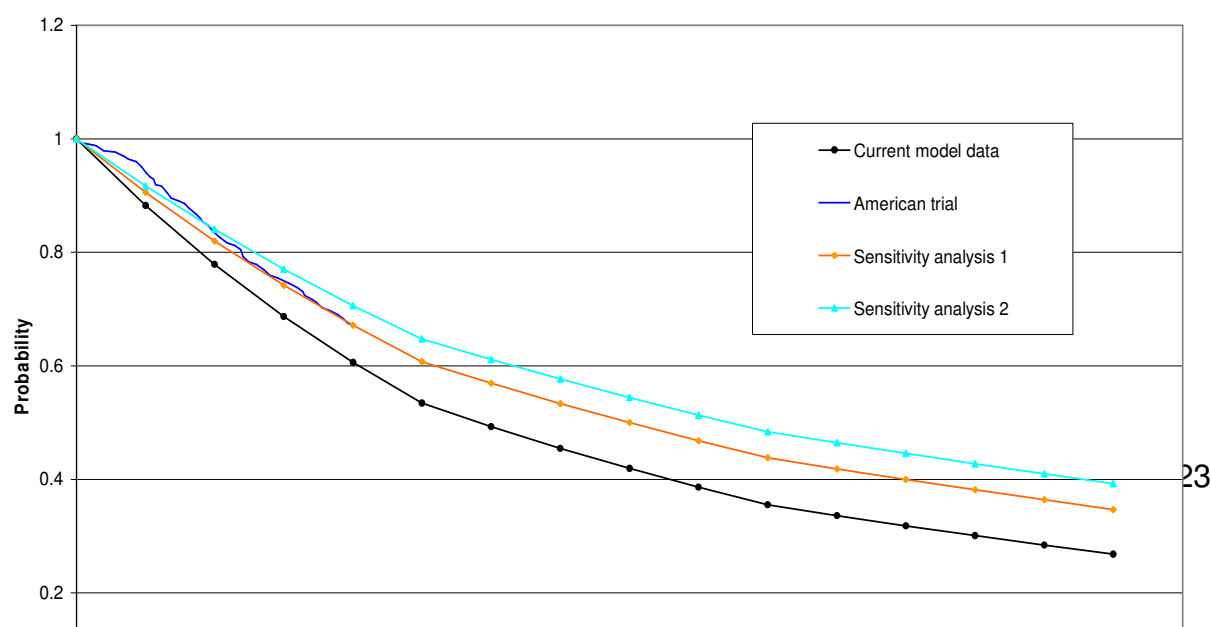
7.2.2.5 Rate of recurrence over time in the comparator arm

Analysis of research papers such as Stal et al⁴³ and Press et al⁴⁴ suggest that there is variation in the recurrence rates of the comparator arm due to the many chemotherapy regimens currently available. Patients in the comparator arm of the American trials had a better prognosis than those in the HERA trial, due mainly to the scheduling of trastuzumab and the use of taxanes. In order to analyse the cost-effectiveness of trastuzumab when compared to a different baseline, the comparator arm from the American trials was applied as closely as was feasible to the Roche model by altering the relative recurrence rates over time. Two scenarios were applied to the ERG's suggested base-case, shown graphically below. The effect on the ICER is shown in Table 24.

Table 24: ERG sensitivity analysis – rate of recurrence in comparator arm

	Relative risk of recurrence over time in the comparator arm			ICER
	1 to 5 years	5 to 10 years	10 to 45 years	
Modelled scenario	1	0.64	0.41	£18,449
Sensitivity analysis 1	0.8	0.51	0.33	£19,937
Sensitivity analysis 2	0.7	0.45	0.29	£21,280

Figure 10: Sensitivity analysis around rate of recurrence over time in DFS



The results show that small variations in the comparator arm should not affect the ICER by more than £3,000 since the rate of recurrence in the comparator arm is unlikely to be less than the second sensitivity analysis above, even where the chemotherapy regimen involves a taxane.

Considering the combined effect of decreasing overall survival (to represent data from other research papers) and reducing the baseline risk of recurrence in the comparator arm (to represent the American trial as closely as possible), as above, increases the ICER to **£19,969**.

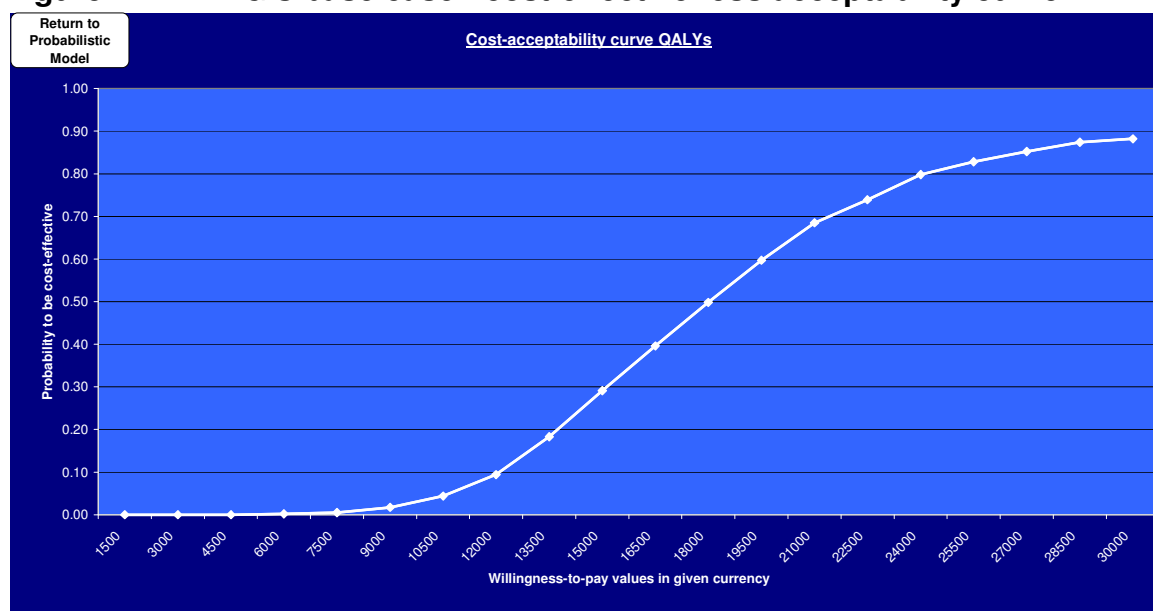
7.2.2.6 Cardiac side effects

Since there is no evidence regarding the long-term cardiac side effects of trastuzumab, the ERG has carried out a sensitivity analysis assuming that the profile of long term cardiac adverse events for trastuzumab will be similar to that seen for anthracyclines. The research paper 'Cardiac Toxicity 4 to 20 Years After Completing Anthracycline Therapy'⁶⁶ suggests that around 23% of patients are likely to have a cardiac adverse event at some time following anthracycline therapy and many who develop late clinical symptoms are likely to die at year 6 – 18. Around half of these did not have a cardiac adverse event during or in the year following therapy. Assuming that 23% of patients experience a cardiac adverse event between years 5 and 20; the ICER increases to **£32,701**. As in Roche's base case model, it has been assumed that none of the patients die from a cardiac adverse event, although the above research paper suggests that this is not the case for anthracycline use. Death as a result of trastuzumab would markedly increase the ICER further.

7.2.3 Probabilistic analysis of ERG's base case

Using the ERG's base case of £18,499 the probabilistic sensitivity analysis from the Rocjhe model produces the following cost-effectiveness acceptability curve.

Figure 11: ERG's base case - cost-effectiveness acceptability curve



The additional analysis carried out by the ERG suggests that the basecase incremental cost-effectiveness ratio should be increased by around £16,000 above the revised basecase of £2,387 presented by Roche. However, it is considered unlikely that the ICER will rise above £20,000 - £30,000 and the cost-acceptability curve above suggests that only at a threshold of less than £20,000 does the probability that trastuzumab is cost-effective fall below 50%.

The ERG has also considered the same base case, but with the assumption that the effect of cardiotoxicity is similar to that of anthracyclines. The results of this probabilistic sensitivity analysis are shown in Figure 12 below.

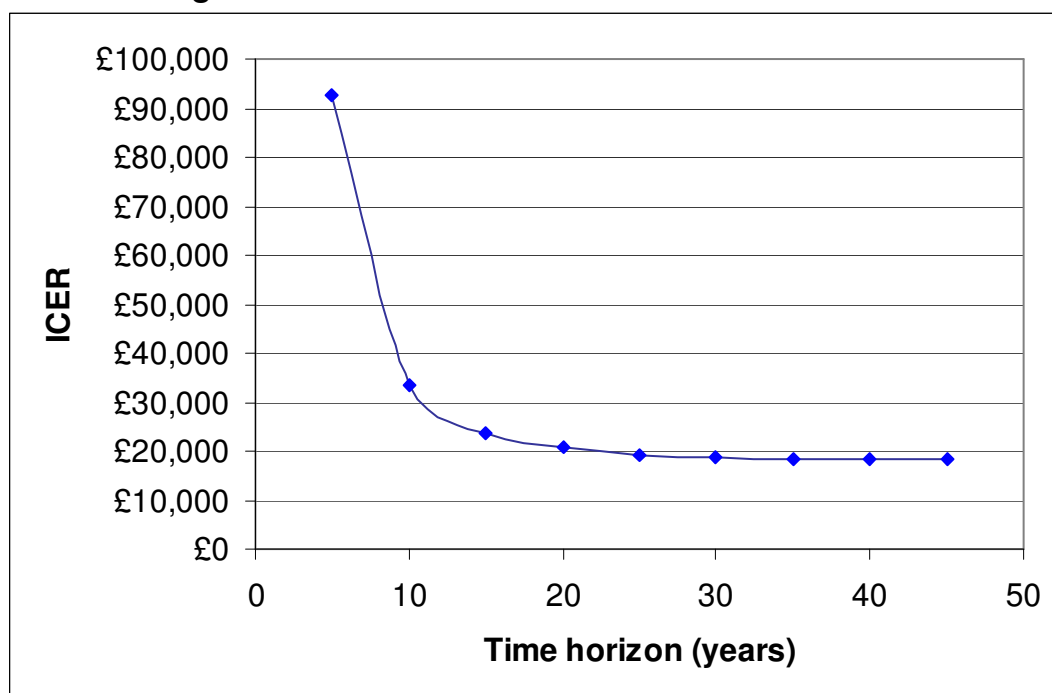
Figure 12: ERG's revised assumption - cost-effectiveness acceptability curve assuming cardiac adverse events as with anthracyclines



7.2.4 Impact of time horizon on model results

Roche's model is based on a lifetime horizon, which follows the guidelines set by NICE. However, analysis of shorter time horizons suggest that the incremental cost per QALY gained of trastuzumab versus no trastuzumab is sensitive to time horizons of up to around 15 years. Beyond this time horizon the economic results are reasonably stable. The impact of different model assumptions regarding parameters and structure are likely to be magnified at time horizons up to 15 years (see figure 13 below).

Figure 13: Impact of time horizon on incremental cost per QALY gained



8 Discussion

8.1 Summary of clinical effectiveness issues

Clinical benefits over a median follow-up of one to two years, especially in the composite outcome, disease-free survival (cancer recurrence or death), are compelling. In the pivotal (HERA) trial and other studies evaluating one year of trastuzumab against observation, the relative hazard of recurrence or death was consistently halved. That these results have been repeated over five trials adds to their plausibility.

Clinically significant harms appear to be relatively rare but, while some have suggested cardiotoxicity is manageable, even reversible, cardiologists remain sceptical that the true effects will be known for many years. They draw an analogy with anthracycline therapies, where late effects are known to manifest themselves 4 to 24 years after treatment.

The early closure of the HERA trial and two American studies means that these studies will never *directly* address the issue of clinical utility in overall survival or provide a proper long-term harm-benefit analysis in the unique population they randomised. The credibility of the results, then, rests on the perceived validity of an analogy drawn with standard cytotoxic and hormonal therapies, classes in which short-term gains in disease-free survival have been demonstrated to translate into a fifteen-year advantage in all-cause mortality.

A trial evaluating nine weeks' of trastuzumab versus observation appears to show a similar clinical effect size with no cardiotoxicity, but a small sample size means that we cannot with any certainty guarantee that the former is not a false positive and the latter a false negative. In summary, the results of this cannot be considered definitive but they generate the hypothesis that nine weeks' can be as effective as 52 weeks' therapy, with greater safety and less inconvenience at 50 percent of the cost.

8.2 Summary of cost effectiveness issues

The revised base case presented by Roche produced an estimated cost per QALY of £2,387. However the Roche base case is based on the assumption that the relative risk of recurrence seen in the HERA trial (median follow-up of one year) will remain the same in the trastuzumab arm for ten years and even beyond 10 years, patients will continue to have a reduced rate of recurrence until death. The reduction in recurrence provided by tamoxifen and anthracycline-based chemotherapy drugs for early breast cancer have been shown to continue for around 4 or 5 years after therapy is complete. In addition Roche have not allowed for patients receiving trastuzumab in the metastatic setting who have previously received the drug for early breast cancer. If it is assumed that there is no further benefit beyond five years and that trastuzumab is given to all HER2 positive patients who have a metastatic recurrence, the ICER increases to £18,449. Variations in the comparator arm to allow for different chemotherapy regimens, including taxanes, do not affect the ICER by more than £3,000.

The ICER could potentially increase to over £30,000 if long-term cardiac adverse events became an issue; these effects are unknown due to the short follow-up of the trials.

It is anticipated that the cost-effectiveness of trastuzumab would improve considerably if trastuzumab were provided for nine weekly infusions instead of the twelve month three weekly infusions presented in the Roche report, however further research is required in this area to ensure that this regimen is equally as effective.

Capacity issues are anticipated for HER2 testing, cardiac monitoring and trastuzumab administration and are not included in the model; however any additional money spent in these areas will improve the services for other drugs as well as trastuzumab.

8.3 Implications for research

Further research would be valuable in the following areas:

- 1) Length and schedule of drug administration regimen;
- 2) Longer term follow-up of relative risk of recurrence;
- 3) Analysis of impact of trastuzumab on overall survival;
- 4) Longer term follow-up of effect of cardiac adverse events;
- 5) Implications for resources in chemotherapy suites, HER2 testing laboratories and cardiology departments.

9 Appendices

Appendix A: Further detail of questions for Roche

1. Increased drug cost including cost of administration

It has been assumed that patients would receive three vials on average, although it is much more likely, on average, that patients will receive four vials than two. Predicted additional cost = £2,400 approx.

Please see response in section 6 (B).

It has been assumed that 16 maintenance doses will be administered, although the NCRI UK Clinical Guidelines states that 17 maintenance doses should be provided. Predicted additional cost = £1,400 approx.

Please see table and cost in section 6 (B).

Cost of administration in chemotherapy lounge. Predicted additional cost = £1,500 approx.

Adopting these alternative assumptions results in a base case cost of around £26,400.

Please see response in section 6 (B) above detailing the assumed cost of drug administration and total drug plus administration costs.

2. Sensitivity analysis around the comparator arm

The HERA trial currently has an average follow-up of 1 year. The shape of the DFS curve for the comparator arm beyond the trial period is based on DFS curves from EBCTCG data for all patients with breast cancer. However, no validation has been undertaken to confirm whether or not this extrapolation provides a reasonable representation of disease free survival without treatment of trastuzumab for newly diagnosed patients with HER 2+ breast cancer in the UK. The impact of changes to this assumption on the cost effectiveness ratio is not considered.

Could you provide validation from sources other than the HERA trial to demonstrate that the comparator arm is reasonable, and run a sensitivity analysis around the comparator arm data, taking into account alternative evidence? Suggestions include (i) using available data from the American trials (ii) using natural history data from previous research papers such as those presented in table 25 below.

Table 25: Examples of research papers giving natural history data for HER2 + women

Author	Paper	Year
Press et al	HER-2/neu Gene Amplification Characterised by FISH: Poor prognosis in Node-Negative Breast Carcinomas	Journal of Clinical Oncology; 15 (8); 2894 - 2904, 1997
Stal et al	ErbB2 status and the benefit from two or five years of adjuvant tamoxifen in postmenopausal early stage breast cancer	Annals of Oncology; 11 : 1545 – 1550, 2000
Charpin et al	c-erbB-2 oncoprotein detected by automated quantitative IHC in breast carcinomas correlates with patients' overall and disease-free survival	British Journal of Cancer; 75 (11): 1667 – 1673, 1997

3. Sensitivity analysis around extrapolation of benefits of trastuzumab beyond the trial period.

The cost-effectiveness ratios will be dependent on the assumptions used relating to the future benefits of trastuzumab beyond the trial period. The risk reduction for year 0-5 in the economic model is based on the trial evidence, for which there is an average one year follow-up currently reported. The risk reduction obtained during the trial period has been assumed to remain constant over the remaining period out to year 5.

For years 5 to 10 the risk reduction values are assumed to be the same as years 0 to 5 and for years 11 onwards the risk reduction values are scaled down by around one third. However no discussion or explanation of the rationale behind the selection of values is given. Given the uncertainty surrounding the risk reduction of trastuzumab beyond the trial period it would be useful if further sensitivity analysis could be undertaken around these values to demonstrate the impact of a range of assumptions on the cost effectiveness ratio, particularly to demonstrate the impact of making more conservative assumptions.

4. Confidence intervals (CIs)

The confidence intervals around the RFS , DDFS and RFM curves, given in sheet "Adj and Risk Red Factor" in the economic model remain constant over time, between years 1 to 5, 5 to 10 and 11 years onwards. It seems more reasonable that the CIs should widen considerably over time and therefore the existing assumptions will be underestimating future uncertainty. It would be useful to consider the impact of assuming that the confidence intervals widen over time on the cost effectiveness ratio.

Appendix B: Potential Errors and/or Inconsistencies in the Model

(a) There are several apparent inconsistencies in the cost calculations. See Table 26 below.

Roche have added comments to the ERG proposed figures within the table below. In some cases the original calculations are correct and simply require some further clarification of methods. However in some instances the HTA group have correctly identified small errors in calculations and Roche has subsequently amended these within the model. Those figures contained within the updated model are highlighted in bold.

Table 26: Potential Inconsistencies in the model calculations

Health state / item being costed	Inconsistency	Figure presented	Figure expected	Roche Comment
One year in the recurrence health state - hospitalisation	Incorrect proportion or monthly cost to give the proportional monthly cost presented	<u>(CIC data removed)</u>	<u>(CIC data removed)</u>	<u>(CIC data removed)</u> <u>(CIC data removed)</u>
One year in the recurrence health state - total	Dividing the total by two to give the cost for each recurrence state gives a different value to that presented	<u>(CIC data removed)</u>	<u>(CIC data removed)</u>	<u>(CIC data removed)</u> <u>(CIC data removed)</u>

Health state / item being costed	Inconsistency	Figure presented	Figure expected	Roche Comment
Monthly cost of endocrine therapy/ chemotherapy for metastatic active treatment	Is the assumption that patients have an equal probability of receiving both treatments?	<u>(CIC data removed)</u>	<u>(CIC data removed)</u>	<u>(CIC data removed)</u>
One year in the metastatic state after 1 st year – total	Assuming 3 months of active treatment, the total is inconsistent	<u>(CIC data removed)</u>	<u>(CIC data removed)</u>	<u>(CIC data removed)</u>

Health state / item being costed	Inconsistency	Figure presented	Figure expected	Roche Comment
All of the remaining costs relate to cardiac adverse events table				
Outpatient appointments – grade II - total	$\pounds 344 + (\pounds 60 \times 3 \times 0.5) + (\pounds 344 \times 2 \times 0.5)$	£606	£778	<p>The figure of £606 is correct:</p> <p>1 initial + 1 FU @ £172 = $\pounds 172 \times 2 = \pounds 344$</p> <p>50% have 3 GP visits/yr @ £60 = $\pounds 60 \times 3 = \pounds 180 \times 50\% = \pounds 90$</p> <p>50% have 2 FU/yr with cardiologist @ £172 = $\pounds 172 \times 2 = \pounds 344 \times 50\% = \pounds 172$</p> <p>Therefore, $\pounds 172 + \pounds 90 + \pounds 344 = \pounds 606$</p>
Days in hospital – grade III – total cost	$\pounds 460 \times 8.1 \text{ days} \times 47.7\% = \pounds 1777.30$ Readmission = 29.6% of £1777.30 = £526.08 □ Total = $\pounds 1777.30 + \pounds 526.08$	£2,768	£2,303	Agreed, new figure included within model.

Health state / item being costed	Inconsistency	Figure presented	Figure expected	Roche Comment
Days in hospital – grade IV	Readmission cost excluded	£3,558	£5056	Agreed, new figure included within model.
Diagnostic tests – grade I	Grade I does not exclude any of the tests given to Grade II and III adverse events, but the cost is lower	£442	£451	Agreed, new figure included within model.
Medication – grade IV – total cost	The sum of each of the medications presented is inconsistent with the total	£1,113	£1,005	<p>The costs we presented of £1,113 are correct:</p> <p>ACE inhibitor: £181.25</p> <p>β blocker: £107.06</p> <p>Diuretic: £10.16</p> <p>Spironolactone: £27.38</p> <p>Digoxin: £51.10</p> <p>Kidney function test: £28</p> <p>5.9% of patients get ICD, plus cost of implant (£12,000*5.9%=£708)</p> <p>Total: £1,112.95</p>

(b) In addition please can you provide a brief explanation/ clarification of the following:

Why a half-cycle correction was only applied to the death state?

This was an error within the writing of the submission. The model actually makes a half-cycle correction to all health states within the model.

The time spent in each health state for the probabilistic model is missing. Was this done in error?

The time spent in each health state for the probabilistic analysis is located in sheet: "CEA probabilistic model", row 61 to row 69.

Data source query - where is the 'Incidence of other Cardiac AE's for trastuzumab only' figure of 1.2 derived from?

1.2% is the percent of the population in the trastuzumab arm that experience cardiac AEs. The 1.2 % is calculated directly from the HERA publication (Piccart-Gebhart 2005, Table 2, p 1665, see also extensive footnotes in table). The table lists severe CHF: 0.54% and symptomatic CHF, including severe CHF as 1.73%. Therefore subtracting severe events from 1.73% provides the 1.2% figure for symptomatic CHF.

Appendix C: Quality Assessment Checklist based on original submission

Issue	Yes Or No	Comments
1. A statement of the problem	Y	Problem clearly defined.
2. A discussion of the need for modeling vs alternatives	N	There is no discussion surrounding any alternatives, although there is some justification for using the type of model employed
3. A description of the relevant factors and outcomes	Y	
4. A description of the model, including reasons for this type of model and a specification of the scope, time frame, perspective, comparators and settings	Y	Roche's scope was combined with a description of the HERA trial results and did not clearly define the population or outcomes of interest. There was no reference to cardiac contraindications or level of fitness in the patient population. There was also no clear description of the dosage and administration of trastuzumab. The model was clearly described, although reasons for the type of model used were not given
5. A description of data sources (including subjective estimates) with a description of the strengths and weaknesses of each source, with reference to a specific hierarchy of evidence	Y	Only the HERA trial was described in detail – there is varying detail provided for the other studies collected from the search of clinical trial evidence. The FinHer trial is excluded completely from any descriptions with no valid explanation. In general there was little or no description of potential strengths and weaknesses of data sources used
6. A list of assumptions pertaining to: the structure of the model (eg factors included, relationships and distributions) and the data	Y	There is no explanation as to why Roche used the distributions that they did (i.e. BetaPert distribution). Roche have not explicitly stated the assumptions made regarding extrapolation of the data from the HERA trial. The assumptions relating to extrapolation of the long term benefits were not clearly identified or discussed within the report
7. A list of parameter values that will be used for the basecase analysis, and a list of the ranges of those values that represent appropriate confidence limits for use in sensitivity analysis	Y	Confidence limits for some parameters were not sufficiently wide to represent the worst/ best possible scenario for some of the parameters.
8. The results derived from applying the model for the basecase	Y	Clearly provided in table.
9. The results of the sensitivity analysis: unidimensional , best/worst case: multidimensional/Monte carlo/parateric);threshold	Y	A one-way sensitivity analysis has been carried out around the majority of the parameters. The combined effect of changes to several of these parameter values would have been useful. One-way sensitivity analysis has not been carried out on a number of the important model parameters (i.e. the trastuzumab relative risk parameter after 5 years)

		A CEAC has been provided.
10. A discussion of how the modelling assumptions might affect the results	N	There is no discussion of this, particularly with regards to the assumptions surrounding extrapolation of the HERA data.
11. A description of the validation undertaken including: concurrence of experts internal consistency, external consistency predictive validity	N	<p>It has been verified that the model is representative of the HERA trial and there is some validation surrounding the MEDTAP studies of cost parameters and health utility scores. However, there is no validation of the disease natural history or validation of the model outputs beyond the first two years.</p> <p>Since 11% of the patients participating in the HERA trial received trastuzumab in the neo-adjuvant setting, there is a small issue with external validity which has not been discussed in the Roche report.</p>
12. A description of the setting to which the results can be applied	Y	There are no guidelines in the report surrounding the health levels required of the patients, particularly with regards to low LVEF scores.
13. A description of research in progress that could yield new data that could alter the results of the analysis	Y	The results of the FinHer study suggest that trastuzumab may be as effective if provided in nine weekly infusions rather than every three weeks over the period of twelve months. This study met the clinical trial evidence search criteria but was not subsequently discussed in detail. There was no discussion of the potential to provide a shorter course of the drug.

10 References

1. Piccart-Gebhart, M. J., Procter, M., Leyland-Jones, B., Goldhirsch, A., Untch, M., Smith, I., Gianni, L., Baselga, J., Bell, R., Jackisch, C., Cameron, D., Dowsett, M., Barrios, C. H., Steger, G., Huang, C. S., Andersson, M., Inbar, M., Lichinitser, M., Lang, I., Nitz, U., Iwata, H., Thomssen, C., Lohrisch, C., Suter, T. M., Ruschoff, J., Suto, T., Grotzer, V., Ward, C., Strahle, C., McFadden, E., Dolci, M. S., Gelber, R. D., and Herceptin Adjuvant (HERA) Trial Study Team. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer.[see comment]. *New England Journal of Medicine* 20-10-2005; **353** 1659-1672.
2. Karnon, J, Delea, T, Johnston, S, Smith, R, and Brandman, J Cost Effectiveness of Extended Adjuvant Letrozole in Postmenopausal Women after Adjuvant Tamoxifen Therapy. *Pharmacoeconomics* 2006;
3. Remak, E and Brazil, L Cost of managing women presenting with stage IV breast cancer in the United Kingdom. *British Journal of Cancer* 2004; **91** 77-83.
4. Joensuu, H., Kellokumpu-Lehtinen, P. L., Bono, P., Alanko, T., Kataja, V., Asola, R., Utriainen, T., Kokko, R., Hemminki, A., Tarkkanen, M., Turpeenniemi-Hujanen, T., Jyrkkio, S., Flander, M., Helle, L., Ingalsuo, S., Johansson, K., Jaaskelainen, A. S., Pajunen, M., Rauhaluoma, M., Kaleva-Kerola, J., Salminen, T., Leinonen, M., Elomaa, I., Isola, J., and FinHer, Study, I Adjuvant docetaxel or vinorelbine with or without trastuzumab for breast cancer.[see comment]. *New England Journal of Medicine* 23-2-2006; **354** 809-820.
5. NCRI Breast Clinical Studies Group UK Clinical Guidelines for the Use of Adjuvant Trastuzumab (Herceptin®) With or Following Chemotherapy in HER2-positive Early Breast Cancer. 2005;
6. Phase III Randomized Study of Trastuzumab (Herceptin®) Versus No Herceptin in Women With HER-2 Positive Primary Breast Cancer Who Have Completed Adjuvant Chemotherapy (www.clinicaltrials.gov). 2002;
7. Phase III Randomized Study of Doxorubicin Plus Cyclophosphamide Followed By Paclitaxel With or Without Trastuzumab (Herceptin®) in Women With HER-2-Overexpressing Node-Positive or High-Risk Node-Negative Breast Cancer (www.clinicaltrials.gov). 2006;
8. Phase III Randomized Study of Doxorubicin and Cyclophosphamide Followed By Paclitaxel With or Without Trastuzumab (Herceptin®) in

Women With Node-Positive Breast Cancer That Overexpresses HER2 (www.clinicaltrials.gov). 2005;

9. Phase III Randomized Study of Adjuvant Doxorubicin, Cyclophosphamide, and Docetaxel With or Without Trastuzumab (Herceptin) Versus Trastuzumab, Docetaxel, and Either Carboplatin or Cisplatin in Women With HER2-neu-Expressing Node-Positive or High-Risk Node-Negative Operable Breast Cancer (www.clinicaltrials.gov). 2006;
10. Trastuzumab In Combination Therapy For Her2-Positive Early Breast Cancer Patients. *San Antonio Breast Cancer Symposium Newsletter* 2005; **2** 2-3.
11. Menard S, Calini P, Tomasic G, and et al Pathologic identification of two distinct breast carcinoma subsets with diverging clinical behaviours. *Breast Cancer Res Treat* 1999; **55** 169-177.
12. Pietras RJ, Pegram MD, Finn RS, Maneval DA, and Slamon DJ Remission of human breast cancer xenografts on therapy with humanized monoclonal antibody to HER2 receptor and DNA-reactive drugs. *Oncogene* 1998; **17** 2235-2249.
13. Early Breast Cancer Trialists' Collaborative Group Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials.[see comment]. *Lancet* 14-5-2005; **365** 1687-1717.
14. Saphner T, Tormey DC, and Gray R Annual hazard rates of recurrence for breast cancer after primary therapy. *Journal of Clinical Oncology* 1996; **14** 2738-2746.
15. National Institute for Clinical Excellence Improving outcomes in breast cancer. Guidance on cancer services. Manual update. 2002;
16. Slamon DJ, Clark GM, and Wong SG Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* 1987; **235** 177-182.
17. Ross JS and Fletcher JA HER2/neu (c-erb-B2) gene and protein in breast cancer. *Am J Clin Pathol* 1999; **112** S53-S67.
18. Chilcott J, Lloyd Jones M, and Wilkinson A Docetaxel for the adjuvant treatment of early node-positive breast cancer: a single technology appraisal. 2006;
19. International Conference on Harmonisation E9 Expert Working Group ICH Harmonised Tripartite Guideline. Statistical principles for clinical trials. *Stat Med* 1999; **18** 1905-1942.
20. Freemantle, Nick, Calvert, Melanie, Wood, John, Eastaugh, Joanne, and Griffin, Carl Composite Outcomes in Randomized Trials: Greater

Precision But With Greater Uncertainty? *JAMA: The Journal of the American Medical Association* 21-5-2003; **289** 2554-2559.

21. Baum, M. Current status of aromatase inhibitors in the management of breast cancer and critique of the NCIC MA-17 trial. *Cancer Control* 2004; **11** 217-221.
22. Fleming, Thomas R. and DeMets, David L. Surrogate End Points in Clinical Trials: Are We Being Misled? *Annals of Internal Medicine* 1-10-1996; **125** 605-613.
23. Guidance for Industry: Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics. *US Food and Drug Administration* 2005;
24. Romond, E. H., Perez, E. A., Bryant, J., Suman, V. J., Geyer, C. E., Jr., Davidson, N. E., Tan-Chiu, E., Martino, S., Paik, S., Kaufman, P. A., Swain, S. M., Pisansky, T. M., Fehrenbacher, L., Kutteh, L. A., Vogel, V. G., Visscher, D. W., Yothers, G., Jenkins, R. B., Brown, A. M., Dakhil, S. R., Mamounas, E. P., Lingle, W. L., Klein, P. M., Ingle, J. N., and Wolmark, N. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer.[see comment]. *New England Journal of Medicine* 20-10-2005; **353** 1673-1684.
25. Abrams, Keith R. Monitoring randomised controlled trials. *BMJ* 18-4-1998; **316** 1183-1184.
26. Pocock, Stuart J. P. When (Not) to Stop a Clinical Trial for Benefit. [Editorial]. *JAMA* 2-11-2005; **294** 2228-2230.
27. Pocock, Stuart and White, Ian Trials stopped early: too good to be true? *The Lancet* 20-3-1999; **353** 943-944.
28. Rothwell PM External validity of randomised controlled trials: "To whom do the results of this trial apply?". *Lancet* 2005; **365** 82-93.
29. Ray, Wayne A. Population-Based Studies of Adverse Drug Effects. [Miscellaneous Article]. *New England Journal of Medicine* 23-10-2003; **349** 1592-1594.
30. Kaufman DW and Shapiro S Epidemiological assessment of drug-induced disease. *Lancet* 2000; **356** 1339-1343.
31. Vandenbroucke, Jan P. Benefits and harms of drug treatments. *BMJ* 3-7-2004; **329** 2-3.
32. Dieppe, Paul, Bartlett, Christopher, Davey, Peter, Doyal, Lesley, and Ebrahim, Shah Balancing benefits and harms: the example of non-steroidal anti-inflammatory drugs. *BMJ* 3-7-2004; **329** 31-34.
33. Slamon D, Eiermann W, Robert N, Pienkowski T, Martin M, Pawlicki M, and et al Phase III Trial Comparing AC-T with AC-TH and with TCH in the Adjuvant Treatment of HER2 positive Early Breast Cancer Patients:

First Interim Efficacy Analysis (Conference Presentation: SABCS 2005). 2006;

34. Schulz KF, Chalmers I, Hayes RJ, and Altman D Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995; **273** 408-412.
35. Wheatley, Keith and Clayton, David Be skeptical about unexpected large apparent treatment effects:: the case of an MRC AML12 randomization. *Controlled Clinical Trials* 2003; **24** 66-70.
36. DCTD, NCI, NIH, and DHHS Cancer Therapy Evaluation Program, Common Terminology Criteria for Adverse Events, Version 3.0. 31-3-2003;
37. The Criteria Committee of the New York Heart Association Nomenclature and Criteria for Diagnosis of Diseases of the Heart and Great Vessels. 1994; **9**
38. Ewer, Michael S., Vooletich, Mary T., Durand, Jean Bernard, Woods, Myrshia L., Davis, Joseph R., Valero, Vicente, and Lenihan, Daniel J. Reversibility of Trastuzumab-Related Cardiotoxicity: New Insights Based on Clinical Course and Response to Medical Treatment. *Journal of Clinical Oncology* 1-11-2005; **23** 7820-7826.
39. Ewer, Michael S. and Lippman, Scott M. Type II Chemotherapy-Related Cardiac Dysfunction: Time to Recognize a New Entity. *Journal of Clinical Oncology* 1-5-2005; **23** 2900-2902.
40. Tan-Chiu, E., Yothers, G., Romond, E., Geyer, C. E., Jr., Ewer, M., Keefe, D., Shannon, R. P., Swain, S. M., Brown, A., Fehrenbacher, L., Vogel, V. G., Seay, T. E., Rastogi, P., Mamounas, E. P., Wolmark, N., and Bryant, J. Assessment of cardiac dysfunction in a randomized trial comparing doxorubicin and cyclophosphamide followed by paclitaxel, with or without trastuzumab as adjuvant therapy in node-positive, human epidermal growth factor receptor 2-overexpressing breast cancer: NSABP B-31.[see comment]. *Journal of Clinical Oncology* 1-11-2005; **23** 7811-7819.
41. Yeh, Edward T. H., Tong, Ann T., Lenihan, Daniel J., Yusuf, S. Wamique, Swafford, Joseph, Champion, Christopher, Durand, Jean Bernard, Gibbs, Harry, Zafarmand, Alireza Atef, and Ewer, Michael S. Cardiovascular Complications of Cancer Therapy: Diagnosis, Pathogenesis, and Management. *Circulation* 29-6-2004; **109** 3122-3131.
42. Slamon, D. J., Leyland-Jones, B., Shak, S., Fuchs, H., Paton, V., Bajamonde, A., Fleming, T., Eiermann, W., Wolter, J., Pegram, M., Baselga, J., and Norton, L. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses

HER2.[see comment]. *New England Journal of Medicine* 15-3-2001; **344** 783-792.

43. Stal, O, Borg, A, Ferno, M, Kallstrom, C, Malmstrom, P, and Nordenskjold, B ErbB2 status and the benefit from two or five years of adjuvant tamoxifen in postmenopausal early stage breast cancer. *Annals of Oncology* 2000;
44. Press, M, Bernstein, L, Thomas, P, and Meisner, L HER-2/neu Gene Amplification Characterized by Fluorescence In Situ Hybridization: Poor Prognosis in Node-Negative Breast Carcinomas. *Journal of Clinical Oncology* 1997;
45. Joensuu, H, Kellokumpu-Lehtinen, P-L, Bono, P, Alano, T, Kataja, V, Asola, R, Utrieinen, T, Kokko, R, Blanco, G, and Isola, J Trastuzumab in combination with docetaxel or vinorelbine as adjuvant treatment of breast cancer: the FinHer Trial. *San Antonio Breast Cancer Symposium 2005* 2005;
46. Moliterni, A, Menard, S, Valagussa, P, Biganzoli, P, and Boracchi, P HER2 Overexpression and Doxorubicin in Adjuvant Chemotherapy for Resectable Breast Cancer. *Journal of Clinical Oncology* 2006; **21** 458-462.
47. Mastro, L, Bruzzi, P, Nicolo, G, Cavazzini, G, and Contu, A HER2 Expression and efficacy of dose-dense anthracycline-containing adjuvant chemotherapy in breast cancer patients. *British Journal of Cancer* 2005; **93** 7-14.
48. Nieto, Y, Nawaz, S, Shpall, E, and Bearman, S Long-Term Analysis and Prospective Validation of a Prognostic Model for Patients with High-Risk Primary Breast Cancer Receiving High-Dose Chemotherapy. *Clinical Cancer Research* 2004;
49. Trastuzumab (Herceptin) for Early Breast Cancer: Thames Valley Discussion Paper for Consultation. 2006;
50. Dowsett, M, Bartlett, J, Ellis, I, Salter, J, and Hills, M Correlation between immunohistochemistry (HercepTest) and fluorescence in situ hybridization (FISH) for HER-2 in 426 breast carcinomas from 37 centres. *Journal of pathology* 2003; **199** 418-423.
51. Yaziji H and Goldstein, L HER-2 Testing in Breast Cancer Using Parallel Tissue-Based Methods. *JAMA* 2004; **291** 1972-1977.
52. Bilous, M, Ades, C, Armes, J, and Bishop, J Predicting the HER2 status of breast cancer from basic histopathology data: an analysis of 1500 breast cancers as part of the HER2000 International Study. *The Breast* 2003;
53. Dr.Andy Hanby 2006;

54. Taken from DoH shadow tariff. 2006;
55. Cox DR Regression models and life tables. *J Roy Statist Soc B* 1972; **34** 187-220.
56. Altman DG Practical Statistics For Medical Research. 1999;
57. Collett D Modelling Survival Data in Medical Research. 1994;
58. Juni, P, Rutjes, A, and Dieppe, P Are selective Cox 2 inhibitors superior to traditional non steroidal anti-inflammatory drugs? *BMJ* 2002; **324** 1287-1288.
59. Boers, M Seminal pharmaceutical trials: maintaining masking in analysis. *Lancet* 2002; **360** 100-101.
60. The tightening grip of big pharma. *Lancet* 2001; **357** 1141-
61. Pich, J, Carne, X, Arnaiz, J, Gomez, B, Trilla, A, and Rodes, J Role of a research ethics committee in follow-up and publication of results. *Lancet* 2003; **361** 1015-1016.
62. Parmar MK, Torri V, and Stewart L Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. *Stat Med* 1998; **17** 2815-2834.
63. Higgins JP, Thompson SG, Deeks JJ, and et al Measuring inconsistency in meta-analyses. *BMJ* 2003; **327** 557-560.
64. Altman DG and Andersen PK Calculating the number needed to treat for trials where the outcome is time to an event. *BMJ* 1999; **319** 1492-1495.
65. Rosenthal R The "file drawer problem" and tolerance for null results. *Psychological Bulletin* 1979; **86** 638-641.
66. Steinherz, L, Steinherz, P, Tan, C, Heller, G, and Murphy, L Cardiac Toxicity 4 to 20 Years After Completing Anthracycline Therapy. *JAMA* 1991; **266**