# Response to the Evidence Review Group Report commissioned by the NIHR HTA programme on behalf of NICE for Lapatinib and trastuzumab in combination with an aromatase inhibitor for the first-line treatment of metastatic hormone receptor positive breast cancer which over-expresses HER2

## GlaxoSmithKline 28 October 2010

GSK welcomes the opportunity to respond to the Assessment Report produced by the Liverpool Reviews & Implementation Group (LRiG), University of Liverpool Assessment Group. It is hoped that the responses given below to the key comments add some clarity to the issues raised and reduce some of uncertainties in the submission.

Our response is structured to reflect the individual sections in the Assessment Report. However the key issues described below are referred to several times throughout Assessment Group report and therefore the section numbers simply reflect the source of the direct quotes.

## KEY ISSUES

### *Issue 1*

Both manufacturers (GSK and Roche) conducted an indirect comparison to compare lapatinib plus an aromatase inhibitor with trastuzumab plus an aromatase inhibitor. However the AG concluded that such an indirect comparison was inappropriate:

### ***1.1.17 Quantity and Quality of research available***

*"In summary, whilst study designs appear appropriate for the comparison of LAP+AI vs AI or TRA+AI vs AI , key differences in the trials led the AG to the conclusion that it would not be appropriate to pool data or make meaningful comparisons, directly or indirectly, across the two completed trials. This decision was primarily based on differences in patient populations – the key factor being the exclusion of patients in whom the disease was considered by the investigator to be rapidly progressing or life threatening as in EGF30008"*

### *Response*

First, within the 30008 protocol, the exclusion criterion specified above by the AG was incorporated as an amendment on 25th May 2004 to be more consistent with the standards of clinical practice, as this patient population would be more appropriate for treatment with chemotherapy rather than with an aromatase inhibitor.

*"A subject was not eligible for inclusion in this study if any of the following criteria applied:*

*….Subjects with extensive symptomatic visceral disease including hepatic involvement and pulmonary lymphangitic spread of tumor, or the disease was considered by the investigator to be rapidly progressing or life threatening."* (Clinical Study Report EGF30008)

Second, the NICE guidance for the treatment of advanced breast cancer (CG81) from September 2009 (an update and replacement for the Technology Appraisal of capecitabine (TA62, May 2003) advises:

*"1.3.2 - Offer chemotherapy as first-line treatment for patients with ER-positive advanced breast cancer whose disease is imminently life-threatening or requires early relief of symptoms because of significant visceral organ involvement, providing they understand and are prepared to accept the toxicity."*

The above recommendation was based on the Guideline Development Group (GDG) consensus which was supported by a Cochrane review first published online on April 22, 2003 (Wilcken 2003) and subsequently reviewed on 30 August 2006

Third, the AG summarise the above guidance as part of the Executive summary in their report:

*Executive summary, Background:*
*"Currently NICE recommends that endocrine therapy (such as tamoxifen [TAM] or an AI) is offered as first-line treatment to the majority of women with ER+ advanced breast cancer. However, providing patients understand and are prepared to accept the toxicity of chemotherapy, this is recommended as first-line treatment when the ER+ MBC is life-threatening or requires early relief of symptoms because of significant visceral organ involvement. Thus in practice, for patients with HR+/HER2+, TRA is commonly given in combination with chemotherapy."*

Fourth, previous chemotherapy for metastatic disease was an exclusion criterion of the TAnDEM trial. This acknowledges that chemotherapy was a first line treatment option for patients with Hormone Receptor positive metastatic breast cancer at the time patients were enrolled in the TAnDEM trial. Thus, given standard practice, it is not unreasonable to assume that clinicians did administer first-line chemotherapy to patients with life threatening disease or extensive symptomatic visceral organ involvement. In this way standard clinical practice was likely to cause a more similar patient population to be enrolled in both EGF30008 and TAnDEM trials than implied by the AG.

In summary GSK believe that it is realistic to expect physicians working at centres involved in the TAnDEM trial to act in line with the best clinical practice at the time of diagnosis or progression. This would lead to patients with HR+/HER 2+ MBC and life threatening disease or extensive symptomatic visceral organ involvement to be considered for a regime containing chemotherapy, ahead of being entered into the TAnDEM trial. As a result GSK believe that conducting an adjusted indirect comparison seems clinically and methodologically plausible.

Finally, even if there were relevant differences between the populations and study designs of the 30008 and TAnDEM trials, and such differences may confound unadjusted "naïve" indirect comparisons, such differences are not in themselves sufficient to confound an adjusted indirect comparison. Nor is it sufficient that the factor(s) that differ across the trials affect the absolute measure of effectiveness for any given arm. Rather, confounding of adjusted indirect comparisons requires that two conditions are met (1) that there are differences between the trials on some factor related to the population or design of the trials and (2) that this factor modifies or interacts with the measure of treatment effectiveness relative to the comparator(s). Thus the differences between the populations of the EGF30008 and TAnDEM trials in disease severity are not sufficient to confound the indirect comparison unless differences in disease severity modify the measure of relative treatment effectiveness. In our analysis, the measure of relative treatment effectiveness was the

hazard ratio. We know of no evidence to suggest that the HRs for OS differ by disease severity in this population.

***Issue 2***

**Summary of cost-effectiveness evidence: section 1.1.42 Cost-effectiveness review**

*"The AG concludes that in HR+/HER2+ women with MBC, LAP+LET compared with LET is not cost effective. Using a time horizon of 20 years, the AG estimates an ICER which exceeds £220,000 per QALY gained for the comparison of LAP+LET vs LET; the incremental total costs and QALYs per patient treated are estimated as £25,209 and 0.114 respectively."*

***Response***

For the comparison of LAP + LET versus LET, the ICER estimated by the AG, of > £220,000 per QALY gained, far exceeds the estimates produced by both manufacturers (GSK and Roche). Whilst there are differences in the base case estimates of total costs, the difference in the ICERs calculated by the manufacturers and that calculated by the AG is primarily driven by differences in the incremental QALY estimates. The ICER calculated by GSK for LAP+LET compared with LET was £74,448 per QALY gained with incremental total costs of £34,737 and an incremental QALY of 0.467 per patient treated. From the data in Table 23 of the AG report the estimate of incremental QALY for this same comparison based on Roche data is 0.42. The AG estimate of incremental QALY for LAP + LET versus LET of 0.114 is thus approximately 3.5 to 4 times lower than that estimated by the manufacturers and is substantially different from the AG estimate for TRA + ANA versus ANA (0.448 QALYs gained). The magnitude of the difference in the AG estimates of efficacy for the two interventions which are a combination of the same class of drugs (an anti-HER2 agent plus an aromatase inhibitor) is unexpected and this data is further discussed below.

The modelling approach taken by the AG differs from that used by GSK to estimate the cost effectiveness of lapatinib plus an aromatase inhibitor. The AG justifies their approach based on a number of theoretical arguments however no evidence justifying the use of the AG modelling approach over the one adopted by GSK has been provided. GSK is unaware of any empirical data that supports the hypothesis that modelling PFS and PPS independently yields more accurate projections of overall survival (OS) than those based on parametric models based on OS data. In the absence of empirical data from secondary sources GSK believe that the best way to assess the models is to compare the output to the clinical trial data.

**Comparative analysis of AG and GSK modelling approaches**

Since the AG did not report survival curves from their model, it was necessary to approximate these values based on the data provided. To do so, a simple Markov model was constructed with daily cycle length and a timeframe of 10 years (3652 days). States of the model were the same as those used by the AG and GSK models (i.e., PFS, PPS, and Dead). The following assumptions were made to approximate the AG modelling approach.

1. Probabilities of PFS events in each month up to 500 days were based on treatment group specific Kaplan Meier survival probabilities for PFS. Thereafter, PFS was assumed to be the same for both groups and was estimated by fitting an exponential model to PFS after 500 days using accelerated failure time regression. Based on this analysis, the Area Under the Curve (AUC) for PFS at 500 days, corresponding to expected PFS at this point, was estimated to be 197.5 days for LET and 265.1 days for LAP+LET (difference=67.6 days). This closely matches the result obtained from

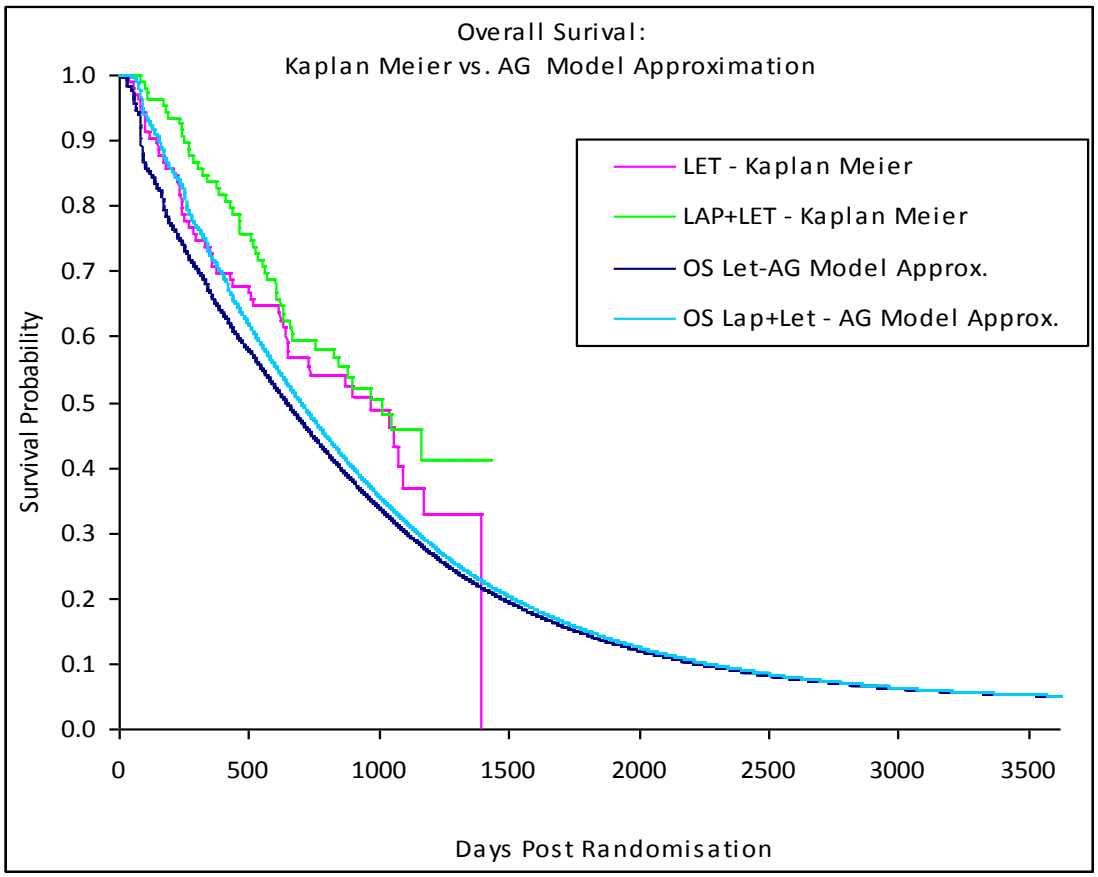
the AG report. The AG did not report their estimates of PFS after 500 days. Based on the fitted exponential model, it was estimated that expected time to death or progression for patients remaining alive and without progression at 500 days is 479.3 days (lambda for exponential model = 0.002086). For each treatment group, approximately 20% of patients remaining alive and progression free at that point thus contribute approximately 80 days to expected survival for the entire cohort over the 10 year timeframe.

2. Because the estimated proportion of PFS events that were deaths that was used by the AG was not reported in the AG report, these proportions (23.41% for LET and 25.09% for LAP+LET) were estimated by calibrating the model to obtain the estimated OS reported by the AG (928.0 days for LET and 982.8 days for LAP+LET).[1]

3. Probabilities of death given progression were estimated assuming an exponential distribution and mean PPS of 764.8 days (lambda=0.001308) consistent with the mean PPS estimated by the AG.

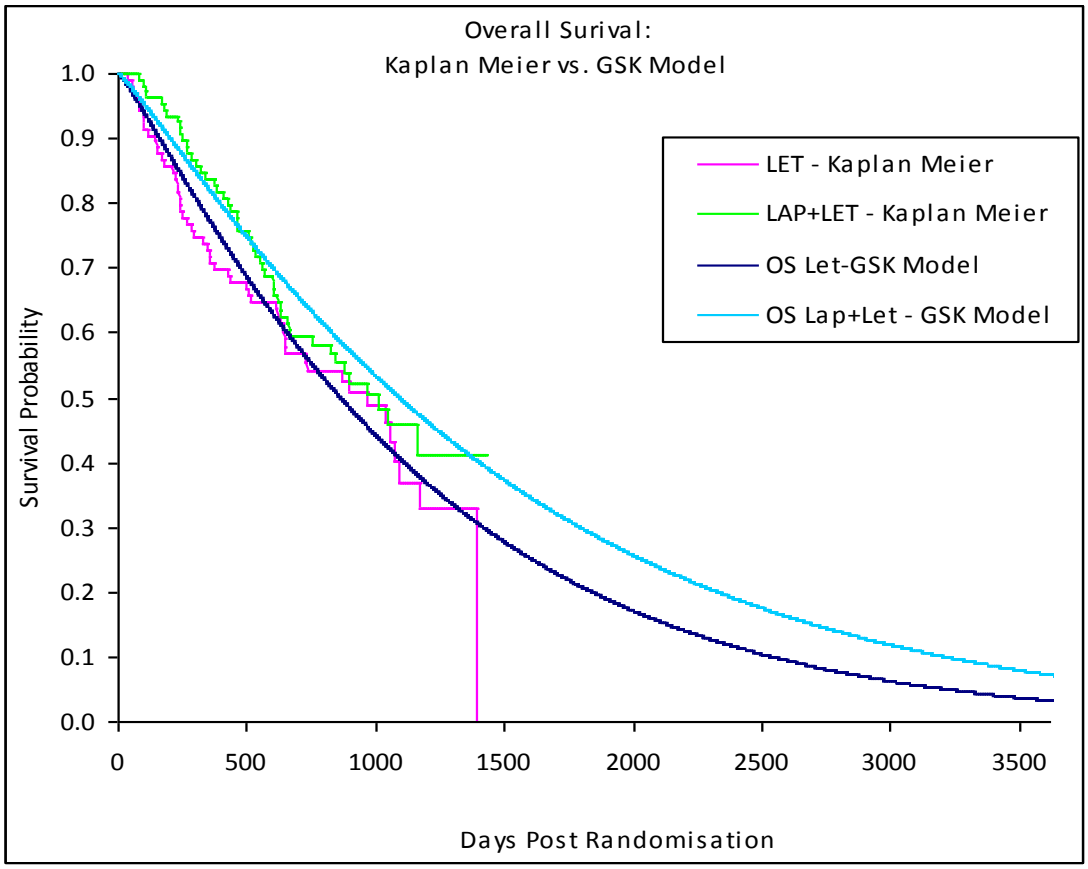
Figure 1 below shows the Kaplan Meier (i.e., empirical) OS curves for LAP+LET and LET from the EGF30008 clinical trial, as well as projections from the AG model approximation. Kaplan-Meier and Weibull model based curves from the GSK model are shown in Figure 2.

**Figure 1. Projected OS: Kaplan-Meier versus AG model approximation**




[1] The actual number of PFS events that were deaths was 21 and 28 in the LET and LAP+LET groups, representing 24% and 32% of all PFS events respectively. The estimates derived by calibration were used to more closely match the results of the AG analysis.

**Figure 2. Projected OS: Kaplan-Meier versus GSK model**




Estimates of expected OS based on each of the projections shown in Figures 1 and 2 are shown in Table 1 below. Results for the Kaplan Meier curves, the AG model approximation, and the GSK model are reported at 1394 days (45.83 months) (the last day for which the empirical OS curve was available for both arms of the trial). Results for the AG model approximation and the GSK model are also reported at 120 months (10 years).

**Table 1. Comparison of area under OS curve based on empirical survival distribution, AG model approximation, and the GSK model**

| Time-frame | Units | Kaplan Meier | | | Models | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | AG Model Approximation | | | GSK | | |
| | | LAP+LET | LET | Diff. | LAP+LET | LET | Diff. | LAP+LET | LET | Diff. |
| 45.83 months (Trial FU) | Years | 2.28 | 2.51 | 0.23 | 2.073 | 1.943 | 0.130 | 2.566 | 2.310 | 0.257 |
| | Days | 833.0 | 917.9 | 84.8 | 756.5 | 709.1 | 47.4 | 936.7 | 843.0 | 93.6 |
| | Difference vs. KM | | | | -9% | -23% | -44% | 2% | 1% | 10% |
| 120 months | Years | na | | | 2.693 | 2.542 | 0.150 | 3.759 | 3.072 | 0.687 |
| | Days | | | | 982.8 | 928.0 | 54.8 | 1372.1 | 1121.4 | 250.7 |

Results from the AG model approximation exactly replicate the expected OS for LAP+LET and LET reported by the AG report (982.8 days for LAP+LET and 928.0 days for LET). The discounted life years (LYs) from the AG model approximation are also similar but slightly less than those reported by the AG (2.517 for LAP+LET and 2.374 for LET from the approximation compared with 2.693 for LAP+LET and 2.549 for LAP from the AG model report). The difference in discounted expected LYs is virtually the same however (0.143 from the AG model approximation and 0.144 from the AG model). Expected LYs from the AG model approximation is approximately 7% less than those from the AG model. The source of this discrepancy is uncertain but may relate to the methods used to discount LYs.

The OS curves based on the AG model approximation are substantially less than the empirical OS curves for both LAP+LET and LET (Figure 1). The extent of the underestimation is greater for LAP+LET (23% underestimate for LAP+LET vs. 9% underestimate for LET). Based on the AG model approximation, the estimated difference in expected OS for LAP + LET vs. LET up to the end of follow-up is 47.4 days. This is 44% less than that based on the Kaplan-Meier curves (84.8 days) It should be noted that the latter may represent a conservative estimate of the benefit of LAP+LET, as it assumes no additional benefits for LAP+LET beyond those observed up to the 46 month follow-up period.

In contrast, the GSK model more closely approximates the empirical OS curve for both LAP+LET and LET (Figure 2). The GSK model overestimates OS by 2% for LAP+LET and by 1% for LET only. Based on the GSK model, the difference between LAP+LET and LET in expected OS up to the end of follow-up is 93.6 days, which is 10% greater than that based on the Kaplan-Meier curves. While it must be acknowledged that there is substantial uncertainty regarding OS during that latter part of follow-up (due to small numbers of subjects at risk) and that OS beyond the end of follow-up is unknown, based on this analysis, it is difficult to understand why the AG approach would provide a more reliable projection of OS than the GSK model. The reasons for the discrepancy between the AG model projections and the empirical survival distributions are uncertain, but may relate to the fact that comparisons of PPS break randomization, that censoring on PPS may be affected by informative censoring (because time from PFS to end of study may be correlated with treatment assignment), and because of correlation of PFS and PPS. This analysis highlights the need for further research to identify the most appropriate methods for modelling OS in economic evaluations of treatment for advanced cancer.

The AG uses the same general approach for modelling TRA+ANA versus ANA as they employed for modelling LAP+LET versus LET. However, their estimates of the gain in PFS with TRA+ANA versus ANA (325 days) is approximately five times the gain in PFS projected for LAP+LET versus LET (68 days). Similarly, the gain in OS with TRA+ANA versus ANA (240 days) is more than four times the gain in PFS projected for LAP+LET versus LET (55 days).

The AG results are in contrast with relatively similar hazard ratios (HR) for PFS for TRA+ANA versus ANA (HR=0.60 95%CI 0.45-0.81) (TAnDEM, Kaufman 2009) compared with that for LAP+LET versus LET (HR =0.65 95%CI 0.47-0.89) (Johnston 2009). It is also in contrast to the relatively larger gain in median PFS for LAP+LET versus LET in EGF30008 (8.2 versus 3.0 = 5.2) compared with TRA+ANA versus ANA in TAnDEM (4.8 versus 2.4 months=2.4 months). While the AG discounted the validity of the indirect comparison, it should be noted that the vast majority of the clinical evidence suggest no material differences in effectiveness of LET versus ANA. Indeed in NICE Technology Appraisal 112 (NICE 2006, section 4.3.3) the committee agreed that:

*"there is insufficient evidence to conclude that any one aromatase inhibitor (used within the licensed indications) or treatment strategy is more clinically effective than another."*

Accordingly, unless there is some difference between the populations of the EGF30008 and TAnDEM clinical studies that would materially interact with the relative effectiveness of anti-HER2 therapy, one would not expect differences in the estimated benefits of TRA+ANA versus ANA and LAP+LET versus LET. To this point our data suggest that the differences in the patient populations between the two trials are relatively minor when differences in reporting strategy are taken into account. The fact that the AG approach yields such differences for the two comparisons raises questions regarding the validity of their entire approach.

## POINTS OF CLARIFICATION

### *Item 1*

***1.1.37 Conclusions of the Assessment Group – Independent economic assessment***

*"Metastases at other sites are broadly comparable, with the exception of bone metastases (14.2% in EGF30008 vs 56.5% in TAnDEM) though this is likely to be an artefact of differing reporting methods."*

***Response***

Indeed this is an artefact of differing reporting methods. Within EGF30008 randomisation was stratified within the bone only (excluding other sites) subgroup, hence why it was reported in this fashion. The percentage of patients with metastatic disease which included bone involvement (i.e. on a similar basis to Roche) is shown in Table 2 (Clinical Study Report EGF30008):

**Table 2. Percentage of patients with metastatic disease including bone involvement in the EGF30008 clinical trial**

| | **LAP + LET** | **Placebo + LET** |
|---|---|---|
| Her 2 positive patients | 58% | 61% |
| ITT population | 61% | 61% |

### *Item 2*

***Executive summary, Background and 1.1.6 Impact of health problem***

*"Currently, LAP is recommended for the first-line treatment of breast cancer in England and Wales in combination with capecitabine in the context of clinical trials for women with advanced or metastatic HER2+ breast cancer."*

***Clarification***

Lapatinib in combination with capecitabine is not indicated for the first-line treatment of breast cancer. Lapatinib has been granted a conditional European licence in combination with capecitabine for patients with advanced or metastatic disease with progression following prior therapy, which must have included anthracyclines and taxanes and therapy with trastuzumab in the metastatic setting.

The NICE (draft) guidance (NICE 2010) that was in place at the time of the AG report was as follows:

*"Lapatinib, in combination with capecitabine, is not recommended for the treatment of women with HER2-expressing, advanced or metastatic breast cancer that has progressed*

*following treatment with anthracyclines, taxanes, and trastuzumab in the metastatic setting, except in the context of clinical trials."*

***Item 3***

***Results, 1.1.1 Assessment of clinical effectiveness***

*"To date, the trials do not show a statistically significant benefit in terms of OS for patients taking LAP+LET vs AI monotherapy or TRA+ANA vs AI monotherapy."*

**Clarification**
The EGF 30008 Overall Survival data in the HER2 positive subset has yet to reach maturity.

***Item 4***

***Results, 1.1.1 Assessment of clinical effectiveness***

**"***Further research may be required into treating MBC in the HR+/HER2+ population who are not TRA (or LAP) naive. In addition, future research should consider adjusting for cross-over a priori."*

***Clarification***
An FDA post-approval commitment study EGF114299 is currently being conducted. UK centres are to be involved.

***Item 5***

***1.1.17 Quantity and quality of research available***

*It is not stated that patients received second-line treatment in EGF30008.*

***Clarification***

Second line treatment was at the investigator's discretion and was implicitly permitted. Information on patient second line treatment was sent to NICE, August 2010 in response to questions from the AG.

***Item 6***

***1.1.23 Summary and critique: clinical effectiveness data***

*"However, it is worth repeating that (i) HR+/HER2+ patients are a sub-group of the EGF30008 trial*."

***Clarification***

The primary endpoint of EGF30008 relates to the HER2 positive subgroup. The calculation of the total study population was based on accruing a sufficient number of patients to adequately power the statistical calculations for the HER2 positive treatment comparison.

**ABBREVIATIONS**

| | | |
|---|---|---|
| AG | = | Assessment Group |
| AI | = | aromatase inhibitor |
| ANA | = | anastrozole |
| ER | = | oestrogen receptor |
| ER+ | = | oestrogen receptor-positive |
| HER2 | = | human epidermal growth factor receptor 2 |
| HER2+ | = | HER2-positive |
| HER2- | = | HER2-negative |
| HR | = | hazard ratio |
| HR+ | = | hormone receptor-positive |
| ICER | = | incremental cost-effectiveness ratio |
| ITT | = | intention to treat |
| LAP | = | lapatinib |
| LET | = | letrozole |
| MBC | = | metastatic breast cancer |
| OS | = | overall survival |
| PFS | = | progression-free survival |
| PPS | = | post-progression survival |
| QALY | = | quality adjusted life year |
| TAM | = | tamoxifen |
| TRA | = | trastuzumab |
| vs | = | versus |

**REFERENCES**

Clinical Study Report EGF30008, GSK Data on File (egf30008-01-body-csr)

Kaufman B, Mackey JR, Clemens MR, et al. Trastuzumab plus anastrozole versus anastrozole alone for the treatment of post-menopausal women with human epidermal growth factor receptor 2-positive, hormone receptor–positive metastatic breast cancer: Results from the randomized phase III TAnDEM Study. J Clin Oncol 2009; 27 (33):5529-37.

Johnston S, Pippen J Jr, Pivot X, et al: Lapatinib combined with letrozole versus letrozole and placebo as first-line therapy for post-menopausal hormone receptor–positive metastatic breast cancer. J Clin Oncol 2009; 23 (33):5538-5546.

National Institute for Clinical Excellence. Advanced breast cancer: diagnosis and treatment. (Clinical Guideline No.81). London: NICE, Feb 2009.
http://www.nice.org.uk/nicemedia/live/11778/43414/43414.pdf [Accessed 27th October 2010]

National Institute for Health and Clinical Excellence (NICE). Hormonal therapies for the adjuvant treatment of early oestrogen-receptor-positive breast cancer. Technology Appraisal TA112. 2006
http://www.nice.org.uk/nicemedia/live/11598/33642/33642.pdf [Accessed 28th October 2010]

National Institute for Health and Clinical Excellence (NICE). Breast cancer (advanced or metastatic) - lapatinib: final appraisal determination, June 2010
http://www.nice.org.uk/nicemedia/live/11731/49197/49197.pdf

Wilcken N, Hornbuckle J, Ghersi D. Chemotherapy alone versus endocrine therapy alone for metastatic breast cancer. Cochrane Database of Systematic Reviews 2003, Issue 2. Art. No.: CD002747. DOI: 10.1002/14651858.CD002747