HEALTH TECHNOLOGY APPRAISAL: NICE Health Technology Appraisal - Assessment Report Lapatanib and trastuzumab (with aromatase inhibitor) for locally advanced or metastatic breast cancer	
TO: NICE	**FROM: NHS Quality Improvement Scotland 28 October 2010**

Comment on Assessment Report
Provided by
[redacted]

1. Several publications have now shown that the true incidence of HER2+ disease in the general population is around 15% , rather than 25% as stated on p18.

A higher percentage of metastatic patients may have HER2+ disease, but all data is prior to the routine use of adjuvant Herecptin. This would reduce the percentage from 30% (p18) to closer to 15%

These figures do not affect cost-effectiveness or cost/QUALY but would impact on the total budget costs.

2. The aim of treatment of metastatic disease is stated in several places to be palliation rather than survival. Only on p23 is there a statement that improving survival is an aim.

For HER2 + disease there is definitely improved survival when using chemotherapy + Trastuzumab compared with Trastuzumab alone. The more frequently stated aim in this report is therefore incorrect.

This is not a trivial point. If clinicians feel that the use of chemotherapy + Trastuzumab may prolong survival then they will in most cases offer this option to patients. If the results of the use of an AI alone are sufficiently poor, as they seem to be consistently in the 3 trials assessed here, then the chemotherapy/Trastuzumab option will be dominant. The use of an AI + either Trastuzumab or Lapitinib would not be considered, unless the option to move onto chemotherapy/Trastuzumab later were possible.

AI + either HER2 inhibitor would only be considered in patients unfit for chemotherapy. In these circumstances the very low manufacturers estimate of 50 patients per year in the UK is possibly accurate.
That figure would undoubtedly increase if there was an option of AI + Lapatinib or Trastuzumab to be followed by chemotherapy + Trastuzumab

on progression, as a lower, or delayed, toxicity alternative, and some patients may escape chemotherapy altogether.

In practice therefore the true comparison is chemotherapy/Trastuzumab versus AI/ Trastuzumab or Lapitinib, with a large proportion of these patients progressing to chemotherapy/Trastuzumab.

This is acknowledged on p26 where it is stated that in practice Trastuzumab is commonly given with chemotherapy (although CMF very rarely and Epirubicin almost never!) .

3. The clinical efficacies are mostly assessed by mean time to progression or progression free survival, and these seem to be the basis for the economic calculations.

Please note that the various survival curves and progression free curves are not normally distributed – there is a noticeable skew to the right. That is there are a small proportion of patients who are on the combination therapy for a prolonged period of time. The trial derived curves as detailed in this assessment are all censored by a relatively short duration of follow-up and these patients are therefore missed and not accounted for in the trial reports or the calculations.

This would certainly have an impact on total budget costs, and would probably tend to reduce the costs per QUALY by a small amount, as the annual costs of the drugs is < cost per QUALY

For example using the presented trial data and assuming normally distributed curves resulted in a major underestimation of the total costs for chemotherapy+ Trastuzumab in the metastatic situation.

Comment on Assessment Report
Provided by
[redacted]

The document looks good and is well-researched and well-presented. It naturally looks at the specific questions addressed by the trials and appears to be suggesting that this treatment with AI and trastzumab is not cost-effective. I would look at it another way. That is, if one has a patient who is trastuzumab naïve, has metastatic cancer, and is ER-positive, could trastuzumab plus AI be an alterative to trastuzumab plus chemotherapy? I would have thought the answer from these trials is "yes", and the use of AI plus trastuzumab is unlikely to be more expensive that the chemotherapy combination which they will be given in its place. I would have thought that AI plus trastuzumab would have been a good combination for elderly or less fit patients who need this sort of treatment and using it is likely

to provide a treatment that is cheaper than the current chemotherapy option and one that has less side effects.

Similar arguments apply to lapatinib, but the situation is more complicated here, as its routine use is not as widespread as it the use of trastuzumab.

Comment on Assessment Report
Provided by:
[redacted]

The main comments I would make with regards to the document -

1. Agree that the addition of LAP or TRA to AI seems to be well tolerated with no major increase in toxicities/SAEs and good safety profiles
2. Agree that addition of LAP or TRA to AI leads to an increase in progression free survival and time to progression but no statistically significant increase in overall survival
3. Agree that addition of LAP or TRA to AI is not cost effective when compared to AI monotherapy
4. Agree that indirect comparison of LAP+AI vs TRA+AI is not possible due to differences in the patient populations within the two main trials (TAnDEM and EGF30008)
5. Agree that the trials being considered in the document were carried out when TRA was not standard treatment for patients with HER2+ breast cancer. Patients presenting now with HER2+ early breast cancer will almost all receive TRA as adjuvant treratment. None (or a very small number) of the patients in these trials would have received TRA previously. The vast majority of the patients we now see with HER2+ metastatic breast cancer will have had TRA as adjuvant treatment. Further trials need to assess if TRA is effective in this setting or whether LAP may be more effective

The document is very thorough and very well put together. There are various areas of repitition in the document. There are several points which I would wish to comment on (not sure if I am supposed to comment on "typographical" type errors but will point a few out) -

- on page 8, 26 and 104 it is mentioned that the combination of AI plus TRA or LET would be suitable for around 50 patients per year. It then mentions at bottom of page 18 that it may be around 500 patients. I think that the 500 patient mark would seem more correct (from my own estimates) and wonder if the 50 is in error.

- Agree with authors when it is mentioned on several occasions (eg pages 14 and 52) that indirect comparisons of different AIs and combinations with LAP/TRA are difficult due to the heterogeneous patient populations in the trials. This makes comparisons unreliable.

- Page 33 Bottom of table 2. Data for eLEcTRA trial states that ANA used where LET 2.5mg daily was used

- The EGF30008 trial is referenced to as EFG30008 at various points in the document

- Page 43 - agree it is important to stress that crossover in TAnDEM trial from ANA to TRA + ANA would impact on findings of the trial

- Pages 45-46 - am unsure what the extensive statistical analysis adds to the results. I am suspicious that such analyses are performed to try to get the result that was being looked for at the outset, especially where the initial results were perhaps not as were expected.

Page 51 - agree that extreme caution needs to be taken in findings in Table 35 + 36 in view of the indirect comparisons being made

Page 56 - agree that patient populations are not sufficiently similar between TAnDEM and EGF30008 to allow comparisons to be made. This is borne out by the differences in PFS and TTP between the 2 trials. A point well made by the authors.

Comment on Assessment Report
Provided by:
[redacted]

This document has a number of flaws, some of which suggest inadequate understanding of the context in which the key trials were conducted, as well as those relating to the likely clinical practice in the UK.

In terms of the H-E modelling, it relies on overall survival data which are likely to be dominated by post-trial treatments, and therefore provide poor indicators of the benefit or otherwise of the addition of anti-HER2 therapy to an aromatase inhibitors. This is particularly concerning in that the availability of further anti-HER2 therapy in both arms of this trial for UK patients could be different than for UK patients (where the use of any anti-HER2 therapy beyond first progression is much less than in many other European countries). I am therefore concerned that the basic approach taken to generating the QALY data is flawed and may (or may not) be a poor indicator of the survival of similar patients in the UK context, and thus an inaccurate estimate of eh QALY gains for this particular technology.

This therefore suggests that the conclusions drawn could be erroneous – a better approach would be to ignore the overall survival data as an unreliable estimate of the effect of the intervention in UK practise, and rather to build a model based on the TTP or PFS data which are not confounded by differences in post-trial treatment access to therapies that may be dependant on national variations in drug re-imbursement policies that don't apply to the UK. Whether or not that would change the overall conclusion I have no idea!

Indeed, what all three trials consistently indicate is that monotherapy with an AI in the ER+HER2+ relapsed breast cancer population is a treatment with limited efficacy for most patients, and perhaps therefore the primary question addressed is wrong, given that many of these patients have the option of chemotherapy+trastuzumab. This is particularly true with the widespread, albeit non-licensed use of weekly paclitaxel which can be given to the vast majority of patients with relapsed ER+HER2+ breast cancer, and certainly anyone meeting the basic inclusion criteria for the two larger studies, EGF30008 and Tandem.

Maybe therefore the key question for UK practise would be to compare the *three* strategies of AI monotherapy, AI+anti-HER2 therapy or chemotherapy with a taxane+anti-HER2 therapy, given that the first and last are already licensed and approved by NICE?

Details

I found this document somewhat disconcerting, due to some basic apparent misunderstandings. This may just reflect on a limited involvement of experienced breast cancer clinicians in this report. This may appear a minor point, but nevertheless does not increase the credibility of the conclusions. For example, in the Introduction, and again later on, the text comments on the fact that in general three chemotherapy agents are used to treat ER+HER2+ advanced breast cancers when discussing the alternative strategy of chemotherapy+Trastuzumab. This is not really the case: the best, and only licensed data for the use of trastuzumab with chemotherapy+TRA is with taxane monotherapy. The use of two or three agents is much more common when giving chemotherapy without tratsuzumab, though single agent taxanes, capecitabine, vinorelbine and anthracyclines are also commonly used.

The report provides an estimate of number of patients eligible for AI+anti-HER2 as around 50 per year. Later on it is indicated that this figure originates from the manufacturers – and whilst the report closely analyse the credibility of many other figures provided by the companies, there is no critical discussion around this figure. It is not clear to me how they arrive at the figure of 50/year which I think could be an underestimate :

To get to 50/year we have to assume perhaps 1% relapse/year of which 25% are eligible for AI+anti-HER2 therapy – this seems very low. In the ATAC trial, 1/3 of whose patients came from the UK and only included post-menopausal women with ER+ early breast cancer, a retrospective analysis of the relapse rates by HER2 status reported suggest that amongst the ER+HER2+ post-menopausal women, the recurrence rate is between 20-25% by 5 years……so an average of 4-5% per year. Thus every year (assuming this population only relapses up to 5 years which is a conservative but reasonable estimate) this would then give the equivalent of 25% of the incidence population presenting with metastatic ER+HER2+ disease each year (one fifth of which were diagnosed in each of the previous 5 years).

UK breast cancer incidence = 45 000 of which perhaps 15% are HER2+ = 6 750 of which around 50% are ER+ = 3 400 per year. It is not easy to estimate the proportion of such patients that are post-menopausal, but 50% would be a conservative estimate and that, combined with the data from the ATAC trial (vide supra), would produce a figure of around 400 post-menopausal women relapsing with ER+HER2+ metastatic breast cancer each year across the UK.

To get to 50/year eligible for AI+anti-Her2 therapy, we would have to assume that almost all such patients now either receive adjuvant Trastuzumab (and thus per license also chemotherapy) and/or were candidates for chemotherapy when they relapse. However, many of these patients would not have been treated with chemotherapy, so would not have got Trastuzumab – and whilst there may be some increase in the use of chemotherapy in this population, in order to give access to Trastuzumab, I am not sure how confident one can be of this estimate of 50 patients per year must be very weak and should be acknowledged in this report.

Discussion of the three key trials – whilst it may not have been stated in the publication that patients in EGF300008 received treatment post progression, this is the norm in the treatment of breast cancer, so the authors would probably not have felt this relevant to mention. Thus in my view it should be assumed that all patients in all three studies would likely have been considered for further therapy upon progression. Indeed, the median and mean post-progression survival data is around 2 years (not that dissimilar to many other phase III studies in advanced breast cancer [1]) and being much longer than the initial PFS would be consistent with most patients having access to several lines of therapy unless proven otherwise. Hence the overall survival data for all three trials will be heavily compounded by treatment post progression, and thus not necessarily the best indicator of any survival advantage of this technology in the UK, where access to multiple lines of therapy, particularly multiple lines of therapy including an anti-HER2 agent, is likely to much less than in many other European countries, where for example, Lapatinib is re-imbursed. It is important to note that none of these trials can mandate any particular post-primary progression therapy, since their primary aim was not the survival, or QALY benefit of the intervention,but only to demonstrate what added efficacy occurs when combining an AI with an anti-HER2 therapy as the first line therapy for these patients. It is also strange, given the likely use of further anti-cancer therapies to patients eligible for this intervention, that the only agent that seems to have been included in the cost of therapy post progression was exemestane. In reality, many patients might be considered for one or more lines of chemotherapy, possibly combined with Trastuzumab (but few for Laptinib as it has not been approved for use by NICE).

Exclusion criteria – the report highlights that there is an apparent difference in the exclusion criteria for Tandem and EGF300008. The latter trial has an exclusion of "*life-threatening visceral disease*", and the report surmises that this is responsible for the differences in median overall survival. Whislt this is possible, this exclusion criteria is subjective, and it is highly likely that given the data already in the public domain when Tandem and Electra were

enrolling, that patients who were candidates for chemotherapy +Herceptin by virtue of having "life-threatening visceral disease" would probably not be enrolled. In particular, the exclusion criteria for the tandem trial excluded patients who had "*clinical disease requiring immediate cytotoxic chemo*…" This to a clinician is not really different from the exclusion criteria in the EGF300008 study of "*life-threatening visceral disease*". Furthermore I think therefore that it is disingenuous to imply in the report that the Tandem trial permitted patients to enrol who were " at imminent risk of death": since the protocol also excluded patients with "*uncontrolled serious intercurrent illness*" or "*severe dyspnoea*". Thus the assumption that the better OS in EGF300008 is due to robust application of the exclusion criteria is potentially insecure and so cross-trial comparisons should not be dismissed on that basis alone. Indeed, given that overall survival is much more dependant on post-progression therapy, the best indicator of whether the populations are different with regard to their potential endocrine sensitivity would be to look at the time to progression in the control arms of the two trials, and here very little difference is evident in the ER+HER2+ populations (3 months in the EGF300008 trial and 2.4 or 3.8 in Tandem dependant on whether one uses the ITT or centrally-confirmed ER+ group and 3.3 months for Electra).

Of interest, the text in this report that describes the data from the Electra trial "Interestingly, significant differences were however reported for differences in TTP between the two cohorts of patients that received LET (median: 15.2 months vs. 3.3 months for HR+/HER2+ MBC vs. ***HR+/HER2-*** MBC respectively; HR=0.71; p=0.03)." is at odds with the data in the public domain (SABCAS 2009 conference poster #4094) which clearly indicates that the TTP in the ***HR+HER2+*** patients treated with LET was 3.3 months, and that for the HR+HER2- patients treated with LET alone it was around 15 months from the graph. Whilst this error might simply be a typing mistake (HER2- rather than HER2+), it is critical to understanding the data from the trial.

Comparison of OS in control arms – the report quotes the ITT, NOT the centrally confirmed ER+ (which it is 29 months) although equivalent data are not available for EGF300008. Given that the clinical implementation under consideration is that of the use of the combination of an anti-HER2 therapy in the patient population of post-menopausal women with metastatic ER+HER2+ breast cancer, it is the ER+HER2+ population of the tandem trial that is surely the most relevant to this appraisal. Of note, a recent review of many first-line endocrine studies reported an average median overall survival of 31 months, with which the results of the control arms of both EGF300008 and Tandem could be considered consistent[1i].

Page 10 – EGF300008 is published – JCO November 2009.

I don't understand the criticism of the Tandem trial that "*The manner in which the protocol is implemented in a clinical trial should be clear to all principal investigators to ensure that the same systems and procedures are in place across all centres to reduce protocol violations. This appeared not to be the case in the TAnDEM trial52 where a few major protocol violations were*

identified causing the exclusion of one patient from full analysis and 15 from the PP analysis of efficacy." The occurrence of major protocol violations is NOT evidence that investigators were unclear about the protocol, and whilst the presence of major violations is always of concern, they can occur due to poor communication, simple error or even occasionally deliberate oversight of an eligibility criterion.

Cross-over. The AG correctly notes that there is no accepted method to deal with this, and this causes major problems for any attempt to deduce the cost/QALY gain with an intervention when using overall survival as the end-point from which to deduce QALY benefits. Given this problem, perhaps it is inappropriate to even consider attempting to model the cost/QALY and alternative strategies could be used to estimate the cost-benefit, as in other diseases where survival data are not available. There are recent examples where NICE have not used trial survival data but have modelled it from other data, including the Denosumab data for osteoporosis.

Furthermore, the AG challenges the three basic assumptions behind the RPSFT:

1) *subjects who cross-over are similar to those who do not with regard to important prognostic factors*
2) *no treatment interaction occurs*
3) *the distribution of subjects who do not experience an event is identical between the randomised treatment groups*

1): however, no evidence is supplied that this assumption is NOT met – the argument made is that only progressing patients can cross-over, but the PFS curve shows that almost 100% of cases appear to progress and so therefore if almost all patients are eligible for cross-over, I fail to understand why it is concluded that this criteria is not met.

2): no evidence is provided that there is a treatment interaction to invalidate this assumption. Indeed, any evidence for treatment interaction probably biases *against* a survival advantage for the research arm, since published data suggest that chemo+trastuzumab is significantly better than chemo alone, so patients crossing over from the control arm to receive trastuzumab who also got chemotherapy (the likely use in many patients) could get MORE benefit from the trastuzumab than those in the research arm.

3); since hardly any patients did not experience an event, there is no evidence provided that this assumption is invalidated.

Health economic data.

Given that no firm conclusions can be drawn about the possible survival advantages of either strategy, surely it would be better to ignore the survival data in the trials as being too unreliable and concentrate on the surrogates. The AG dismisses, for example, indirect comparisons between the two studies on the basis of differing outcomes in the control arms – but the

reliability of the overall survival data from the trials as an estimate of what would happen in UK practise is surely at least as uncertain as any cross-trial comparison?

The AG uses PFS + PPS to derive the estimated difference in mean survival – but since the overall survival is driven by the PPS and the treatments administered during that two year period are uncontrolled and unavailable (in particular we have no idea what proportion of patients in either arm received any number of lines of further hormone therapy, chemotherapy or anti_HER2 therapy), how can we have ***any*** confidence that this provides a realistic estimate of the difference in overall survival that would be seen for each of the strategies in a UK context? Surely therefore the approach is fundamentally flawed and would be better done by considering the differences in PFS and response rates as the best indicators of the benefits of allowing the strategy of anti-HER2 + AI. If using these two markers of clinical benefit deems the use of combined AI+anti-HER2 therapy to not be cost effective I think that would be a more credible approach, since the option of crossing over will not be available to UK patients (unlike in the studies which were run internationally where availability of trastuzumab and lapatinib may well have been greater than in the UK).

Disease assessment. I find the text unduly negative – of course investigator assessments may be biased, but as acknowledged, EGF300008 was double blind and Tandem used an independent review committee as it would have been ethically difficult to run the trial as a double-blind study. However, the issue here is that this report is trying to estimate the cost-effectiveness of a strategy, not just the efficacy of a particular treatment, and since the strategy in question is the use of AI+anti-HER2 therapy in routine practice where assessment of progression is by clinicians, the quality of the assessment of progression in the studies is higher than what would happen in daily practice. Furthermore, a similar bias exists in Tandem with regard to toxicity reporting, since most of the toxicities that are higher in the AI+trastuzumab arm are patient reported (arthralgia, headache, diarrhoea, dyspnoea, nausea and vomiting etc.). However, the report does not comment on this source of subjective bias from patients who were necessarily unblended as to which arm of the study they are in.

Electra trial – whilst it is a small and heavily underpowered study, it is interesting to note that the median TTP on the control (AI alone) arm for the patients with ER+HER2+ cancers is very similar to that in the other two studies. This does suggest that at least as an estimate of the benefit of AI monotherapy for this patient group, all three studies provide highly consistent data and this does not appear to be commented upon. It also strongly suggests that AI monotherapy is a relatively ineffective therapy in this particular patient group!

Meta-analysis – I am surprised that the AG did not conduct a meta-analysis of the data from the three trials under consideration. All three test the same question – the potential benefit of combining a non-steroidal AI with anti-

HER2 therapy. The data available do not suggest any reason to think that Trastuzumab and Lapatinib are likely to have very different effects in the population under study, and the control arm outcomes seem very similar suggesting broadly similar populations at least with regard to their potential sensitivity to the hormonal therapy. I would therefore have thought this was an ideal opportunity to conduct such a meta-analysis for efficacy to sharpen the estimates of benefits. Meanwhile, the report is highly critical of the manufacturers' meta-analyses stating that:

1) Crucially, it was unknown if patients with HR+/HER2+ MBC were included in the trials and if so,
2) how many patients were in fact HR+/HER2+…..
3) ….and in many instances, patients who had advanced breast cancer.

With regard to these, it is certain that there will have been ER+HER2+ patients enrolled in the endocrine comparison trials, as acknowledged, but the second observation is also true, that the number is uncertain. However, since the trials in question were randomised and largely conducted before Trastuzumab was available, any knowledge of HER2 status is not likely to have influenced enrolment, and certainly could not have influenced treatment allocation. Thus there will have been HER2+ patients in these studies and they will be balanced between arms, and quite probably across studies.

The third comment again illustrates the lack of knowledge about breast cancer patients. The patients in all the studies being considered, both these endocrine comparison studies that were meta-analysed by the companies, AND the three (tandem, Electra and EGF300008) are advanced breast cancer patients. The use of the terms advanced and metastatic is largely interchangeable, although they are not synonymous – the difference being that advanced also includes locally advanced (stage III) which tends to make up a small proportion of the patients included (5% for example in EGF300008). So this concern is to my mind spurious.

This is not to disagree with the statement that data arising from these indirect comparisons should not be treated with caution – but perhaps the difference in quality of these comparisons with some of the other approaches taken, including using the overall survival data from the three key trials, is perhaps less than is acknowledged in the report.

Response rates – it is stated on page that the only significant difference in response rates was for the tandem trial – but in the ER+Her2+ population in EGF300008 the response rates were 15% and 28% ($p = 0.021$), so I fail to see why this is not considered significant by the AG. Indeed the difference in clinical benefit rate (a clinical end-point which in unselected ER+ patients is associated with improved overall survival) is significant at the 1% level ($p = 0.003$). When these data are first discussed it is commented that the difference is not significant in the ITT population, but of course this includes all patients irrespective of HER2 status, and is therefore an irrelevant data point

for this assessment report, since the strategy under scrutiny is the use of AI+anti-Her2 therapy ***only*** in the ER+HER2+ population.

28 October 2010

[i] Saad et al, Overall Survival and Post-Progression Survival in Advanced Breast Cancer: A Review of Recent Randomized Clinical Trials *J Clin Oncol* 28:1958-1962.