Pharma Mar

Gwpo Zeltia

11 February 2015

[redacted]
Vice chair
National Institute for Health and Care Excellence
10 Spring Gardens
London SW1A 2BU

Dear [redacted].

Re: Final Appraisal Determination – Ovarian cancer - topotecan, pegylated liposomal doxorubicin hydrochloride, paclitaxel, trabectedin and gemcitabine for advanced recurrent disease only (Review ofTA 91 & TA 222) (10468]

Thank you for your letter dated 27 January 2015, setting out your preliminary views with respect *to* the admissibility of the points of appeal raised in our appeal letter dated 20 January 2015. We now respond to these matters. providing additional clarification of the points raised before you make your final decision as to whether each appeal point should be referred on to the Appeal Panel. Our comments are set out below. by reference to the individual appeal point to which they relate.

We also note that you stress that an appeal is not a re-run of the appraisal. We recognise this and would like to emphasise that we do not wish to debate the merits of statistical methods considered during the appraisal. Instead, we wish to highlight that having specifically followed guidelines established by the NICE DSU and earlier adopted by the Evidence Review Group (ERG) of the previous NICE assessment for trabectedin, we consider it unreasonable that important evidence. which by the Assessment Group's own admission would lead to more accurate conclusions for the comparison in question if used, did not form the basis of the recommendation for trabectedin.

1.1 Exclusion by the Appraisal Committee of relevant covariates in the adjusted analysis of trabectedin is unjustified.

The introduction to your letter dated 27 January indicates that your initial scrutiny process is to determine whether any of the points raised by PharmaMar "arguably fall within any one of the grounds" permitted by NICE's appeal process. In the light of that indication, PharmaMar is entitled to assume that you will consider all the information provided in order to establish whether it may arguably fall within any of the permitted grounds of appeal rather than adopt a restrictive assessment of that information based on the numbering or terminology adopted. Therefore, we would have expected that the unreasonableness of this point would have been

considered, given our reference to that concern, rather than a narrow assessment of procedural unfairness.

With this in mind, we would therefore ask that thought is given as to what the ERG of TA222 would have done for this assessment, in considering the extent to which this claim is unreasonable. Two clear facts should be considered when making this judgement:

1. The clinical evidence base has not changed between TA222 and this assessment, other than the interim overall survival results for the trabectedin pivotal study becoming final overall survival results

2. The potential comparators to which trabectedin can be compared has not changed between TA222 and this assessment, indeed the Assessment Group agreed that only topetecan, paclitaxel and PLDH are available for comparison with trabectedin

Hence, there is no evidence to suggest the ERG of TA222 would not reach the same conclusions regarding the methods for appraising trabectedin in this assessment, namely:

a. Evidence synthesis is not required and a direct comparison of trabectedin plus PLDH with PLDH alone is sufficient to address the decision problem (ERG Report TA222)

b. NICE DSU guidance should be followed, and with the availability of final patient level data for this comparison, imbalances in important, prognostic baseline characteristics should be adjusted for, which as discovered in the final analysis includes PFI (Latimer 2013)

This was the expectation of PharmaMar, having already learnt from methodological mistakes made and subsequently corrected by the ERG of TA222. No reasonable decision maker would have come to the current recommendation had it given proper weight to all the relevant guidance and evidence before it, based on the available information.

The point being made here is to show that the course of action taken by the Assessment Group in this assessment and accepted by the Appraisal Committee did not properly take account of guidance in NICE DSU guidelines as well as methods established by the ERG of TA222. This was done without reasonable justification because the final comparator for which the cost-effectiveness of trabectedin plus PLDH was assessed against was still PLDH alone in this assessment (this is because PLDH remained the most cost-effective treatment compared to topetecan and paclitaxel; see Table 142 of TAG Report). However, this comparison was inherently biased in favour of PLDH as it lacked the use of pivotal patient level data and adjustment for imbalances, which skewed results.

We think that it is unreasonable that such judgments and inconsistency could occur in a scientific process such as a technology appraisal and be deemed reasonable. As such, we would like this point also considered as a complaint of unreasonableness under Ground 2.

In addition you have highlighted that this point received considerable discussion in the FAD (4.3.15-4.3.20) and that it may be reasonable that people could differ on these judgments. We would disagree with this conclusion as there is a procedural unfairness in that departing from guidance outlined by the NICE DSU was contrary to PharmaMar's legitimate expectation that NICE would conduct its appraisal in a manner that is consistent with that published guidance and established methods from TA222 {discussed above).

NICE acknowledge that adjusting for characteristics is recommended by the DSU (4.3.16) and concluded that this guidance was not relevant because evidence synthesis was required:

"The focus is on situations in which patient-level data are available, and where evidence synthesis between trials is not required" (Latimer, 2012)

However, as already concluded by the previous ERG of TA222, evidence synthesis was in fact not required for the evaluation of trabectedin {although it may have been for other comparisons not including trabectedin). Even if evidence synthesis were required, the final comparator of interest using evidence synthesis as per the Assessment Group method was ultimately PLDH alone (see Table 142 of TAG report) rendering a direct comparison with adjustment far more appropriate and accurate. In light of this, NICE should have followed DSU guidance and reverted to a fairer direct comparison using patient level data and making appropriate adjustments for imbalances, but neglected to do this. As such we reiterate that this claim still falls under a complaint of fairness under Ground 1a.

1.2. Different interpretation of the evidence by the same Appraisal Committee for the **MTA** and TA222 regarding the use of direct head-to head data for trabectedin to address the decision problem for the non-platinum network (Network 2) is irrational and unfair.

As with 1.1, we note that you believe our complaint may impact on the reasonableness of the Appraisal Committee's recommendations and would arguably also fall within Point 2 of the permitted grounds for appeal. As such, we would like this point also considered as a complaint of unreasonableness under Ground 2.

You state that NICE has repeatedly made the point that one appeal panel does not bind another, however it is highly likely that this panel will only judge this guidance to be unreasonable on the grounds of inconsistency, provided the inconsistency is "very clear

indeed". We feel that in our case, the inconsistency is very clear indeed, and certainly clear enough for this to be considered under Ground 2.

While PharmaMar accepts the position that one Appraisal Committee does not entirely fetter the discretion of a subsequent Committee to depart from its approach where appropriate we disagree that this Appraisal Committee had legitimate reasons to depart from the methods used in TA222 for trabectedin.

As discussed in 1.1:

1. The clinical evidence base has not changed between TA222 and this assessment, other than the interim overall survival results for the trabectedin pivotal study becoming final overall survival results

2. The potential comparators to which trabectedin can be compared has not changed between TA222 and this assessment, indeed the Assessment Group agreed that only topetecan. paclitaxel and PLDH are available for comparison with trabectedin

All in all, the decision problem had not changed for the evaluation of trabectedin within its licensed indication. The same network of trials and same comparators were used in TA222 and this assessment. The only differences in evidence for trabectedin and its comparators were that between the two assessments (STA 2011 and MTA 2014) trabectedin now had final results available and an important imbalance in Platinum Free Interval (PFI) had been discovered in the treatment arms of the trial. However, unreasonably for the evaluation of trabectedin, a complete change in methodology and approach was adopted in this assessment of trabectedin, which did not consider patient level data, or the adjustment of characteristics as did TA222. Although this approach seemed sensible for other comparisons, the choice of using this approach was not justified for trabectedin when considering that PLDH was found to be the most relevant comparator and direct head to head data exists.

This is a clear inconsistency, which in turn led to cost-effectiveness estimates that found trabectedin plus PLDH not to be cost-effective. Were a consistent approach in methods adopted then trabectedin plus PLDH would likely be found cost-effective and recommended for use.

Therefore it was both unreasonable and procedurally unfair for the Appraisal Committee to depart from previous evaluation practice contravening PharmaMar's legitimate expectations that the evidence would be evaluated using the same methods outlined in TA222 (aligned with NICE DSU guidance).

2.1 The Appraisal Committee's rationale for not using adjusted clinical effectiveness results for the cost-effectiveness evaluation of trabectedin in the MTA is flawed and inconsistent with the previous TA222 appraisal and NICE Decision Support Unit guidance.

In your reply you state that this complaint is one of reasonableness rather than procedural unfairness. Given the points made above in relation to the dual nature of the complaint information we have provided we would have anticipated that this point would arguably fall to be considered under both permitted grounds of appeal.

As previously discussed in 1.2, and reiterated here, the scope for evaluating trabectedin did not change. The same clinical trial base was used with the same comparators as per TA222. Therefore, it is unreasonable to change the methods used to evaluate trabectedin unless a plausible rationale is provided.

In this case. the rationale appears to be that in the absence of a consistent dataset for all comparisons, the Assessment Group did not consider it appropriate to analyse a blend of unadjusted and adjusted hazard ratios (4.1.5). We infer from this that by all comparisons, it is meant using unadjusted hazard ratios for trabectedin plus versus paclitaxel and trabectedin plus versus topetecan whilst using adjusted hazard ratios for trabectedin plus PLDH versus PLDH alone. We assume that methods used for other comparisons in other networks outside of trabectedin's marketing authorisation did not affect the methods for evaluating trabectedin, otherwise this introduces further claims for unfairness of the MTA versus STA process.

The rationale provided by the Assessment Group and accepted by the Appraisal Committee is irrational since TA222 concluded that "the relative cost-effectiveness of trabectedin plus PLDH compared to paclitaxel or topotecan monotherapy is not needed" (ERG Report TA222). Given there is no change in the evidence base for these interventions, and the Assessment Group's own model found PLDH alone to be the relevant comparator for which to evaluate cost-effectiveness versus trabectedin plus PLDH, the rationale for applying this new approach to trabectedin is unsubstantiated and unreasonable. In addition from a procedural fairness perspective no notice was provided by the Appraisal Committee that it intended to depart from previous practice or guidance.

We would accept a new approach as reasonable from the approach adopted in TA222 were this justified sufficiently, but this is not the case here. By the Assessment Group's own admission when considering the only relevant comparison (trabectedin plus PLDH versus PLDH alone): "adjustment of clinical effectiveness data for key prognostic factors was likely to result in more accurate estimates of progression-free survival and overall survival". (4.2.19) Therefore. it hardly seems reasonable to ignore the recommendation in

TA222 that evidence synthesis is not required for the evaluation of trabectedin and subsequently adopt a less accurate approach to ensure 'a consistent dataset' whereby the final comparison subsequently excludes the comparators (paclitaxel and topotecan) for which consistency was purposefully sought.

2.2 The Appraisal Committee failed to take into account key differences in baseline characteristics and trial design of relevant studies that have informed the clinical and cost-effectiveness results and subsequent recommendations for the FAD including that of trabectedin.

In the context of the wider MTA these differences in baseline characteristics have been ignored, although it forms the crux of what we believe to be the inappropriate inclusion of some studies into the network meta-analysis (NMA). This NMA forms the basis of trying to evaluate the clinical evidence for ovarian cancer treatments in this appraisal and subsequently is used in the economic analysis. It therefore could lead to very different results and conclusions. We believe the approach taken for inclusion of some of the trials in the NMA is flawed and has ignored expert advice cautioning on its use at the Appraisal Committee Meeting and in responses to the Appraisal Consultation document. Eminent statisticians state that in direct and indirect analysis of evidence, well-informed clinical judgment is required to determine which trials should be combined in such an analysis (Lu and Ades, 2004).

In addition, we do not see anything that shows the Assessment Group has done any sensitivity analysis to see what the results could be were these trials to be excluded from the analysis and how they might influence the results. Comprehensive sensitivity analyses have been done for the economic evaluation but not for the clinical effectiveness results.

The NICE Guide to the methods of technology appraisal (2013) regarding mixed-treatment comparisons (point 5.2.14) states that:

> "If there is doubt about the relevance of a particular trial or set of trials, sensitivity analysis should be presented in which these trials are excluded (or if absent from the base-case analysis, included)."

It would be reasonable to expect, when it is known that such baseline characteristics and trial entry criteria can influence the results as discussed in the FAD (4.1.5, 4.2.9, 4.3.19), that a sensitivity analysis would at least have been undertaken to check their impact – a prime example of where this should have been done is for the Gonzalez-Martin and Parmar trials in network 1 (an explanation of why these differences could influence the results in the NMA were provided in Appendix 2 of our letter dated 20 January). Point 2.5 below

provides a similar example, which you have agreed with, where a failure to conduct appropriate sensitivity analyses on factors that could produce significantly different results and thus influence recommendations.

Data did not exist in the network of evidence including trabectedin to allow consistent use of adjusted comparisons of trabectedin plus PLDH. Namely, paclitaxel and topotecan could not be analysed using patient level data adjusted for imbalances. However, as previously stressed, the Assessment Group of TA222 found that "the relative cost-effectiveness of trabectedin plus PLDH compared to paclitaxel or topotecan monotherapy is not needed" and a direct comparison of trabectedin plus PLDH versus PLDH alone was sufficient to address the decision problem (Assessment Group Report TA222).

Therefore, we challenge the concept that consistency was required in this assessment of trabectedin, whereas previously with the same evidence base and comparators in question, it was deemed that a direct comparison of trabectedin plus PLDH was sufficient. We do not challenge however that this approach should not be used for other networks or comparisons not including trabectedin.

This challenge is further supported by the Assessment Group's final conclusions; the cost-effectiveness estimate of trabectedin £77,000, considered as the most appropriate estimate to use by the Appraisal Committee, is in fact the cost-effectiveness of trabectedin plus PLDH versus PLDH alone. The Assessment Group has consequently shown, using a biased, inaccurate, and unreasonable method, what the Assessment Group of TA222 had already advised: "the relative cost-effectiveness of trabectedin plus PLDH compared to paclitaxel or topotecan monotherapy is not needed".

To highlight the somewhat selective nature of the evidence analysis in this MTA, it should be noted that the quality of life and utility values used in the economic model come from individual patient level data (IPD) analysis from PharmaMar's trial which was conducted for TA2.22 as it was the largest data sample and therefore provided the most robust evidence to fit into health states required for the economic model for all 3 networks. Furthermore, having access to the results of the IPD dataset and Clinical Study Reports from PharmaMar enabled the Assessment Group to have the most robust dataset to model survival curves in network 2 and network 3. So the Assessment Group and Appraisal Committee have happily taken the IPD from PharmaMar for certain analyses as they recognise that this would yield a better and more precise analysis but not analysis of the PharmaMar data in relation to point 1.1, 1.2, 2.1 and 2.3. We see this as unreasonable.

In summary, it is unreasonable that despite the recognition of the importance of differences in baseline characteristics, there has been a failure to take into account the expert feedback and conduct sufficient sensitivity analyses as described in the NICE Methods Guide to determine how the results of the NMA could be influenced. In the case of trabectedin it is unreasonable not to allow the availability of the patient level data to be used to account for the difference in baseline characteristics and accurately represent the clinical effectiveness of the drug.

2.3 The different interpretation of the evidence by the same Appraisal Committee for the MTA and TA222 regarding the use of direct head-to head data for trabectedin to address the decision problem for the non-platinum network (Network 2) is irrational and unfair.

We would repeat the arguments made in Point 2.1 in setting out our view that this complaint arguably falls within both the appeal grounds required to pursue a valid unfairness and unreasonableness appeal.

2.4 An incorrect adjustment by the Assessment Group of drug costs for trabectedin and PLDH has been applied resulting in an inaccurate ICER being calculated.

Thank you for confirming that you are minded to regard this as a valid point of appeal.

2.5 Recommendations within the FAD for the use of paclitaxel within its marketing authorisation are based on extrapolated off-label data and costs in the monotherapy platinum resistant/ refractory patients.

Thank you for confirming that you are mined to agree that this is a valid point of appeal. We also draw your attention that the same rationale for accepting this point as valid should be applied for accepting Point 2.2 above as a valid point.

2.6 Recommendations for use of off-label PLDH in combination with platinum are unlawful.

You state that this is a ground 1b) point and should not be under ground 2. We see that we have placed this under the wrong ground and would like this point to be considered under ground 1b).

You point out that the appraisal has been allowed to proceed to completion and question as to why PharmaMar has not raised this objection previously. We could not do this because the Appraisal Consultation document issued in September 2013 did not contain this recommendation. During 2014 we understand that NICE requested the Department of Health

to enable them to make an off-label recommendation for PLDH in combination with carboplatin. That NICE sought this direction after the consultation and PharmaMar were not notified that it was being sought meant the lawfulness or otherwise of the decision could not have been challenged at that point or indeed until the outcome of the process. As a material consideration this has accordingly prejudiced PharmaMar's position.

We would also expect that, were a recommendation made for the use of paclitaxel in platinum patients based on extrapolated off-label data (as in point 2.5 above), then a similar request would need to be made by NICE to the Department of Health.

In summary, we have demonstrated that the appeal points for which you were minded to say are not valid for consultation with the Appeals Committee have a clear, factually-based claim to be considered as valid and presented to the Appeals Committee. **We reiterate our wish to proceed to an oral hearing**. PharmaMar remains committed to NICE's own internal appeal process but must reserve its rights to pursue judicial review if the outcome of that process does not satisfactorily address its concerns about the lawfulness or otherwise of the previous appraisal process.

Yours sincerely,

[redacted]
Managing Director

References

Lu G and Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Statist. Med. 2004; 23:3105–3124

Latimer L. NICE Decision Support Unit (DSU) Technical Support Document 14: Survival analysis for economic evaluations alongside clinical trials: extrapolation with patient-level data. June 2011 {last updated March 2013).