Issue 1 The TAG requests clarification regarding the details and reasons for censoring within OVA-301

Description of problem	Description of proposed amendment	Justification for amendment
[redacted]	PharmaMar requests that the TAG acknowledges the clarification provided regarding censoring within OVA-301, and consequently it can be concluded that censoring has not unduly biased the interpretation of the results.	We expect that the suggested amendments will provide insight into the reasons for censoring, and give evidence that the survival results of the OVA-301 study are robust.

[redacted]		

[redacted]

[redacted]

Table 10. Reasons of censoring (progression-free survival) in OVA-301.

Reason	Treatment arm				Total	
	PLD		Trabectedin + PLD			
	n	%	n	%	n	%
Censored on the last evaluation before subsequent therapy	88	71.5	90	67.7	178	69.5
Censored on the last tumour assessment without progression (still under evaluation, progression by IA and no further evaluations, lost to follow-up, etc)	19	15.4	29	21.8	48	18.8
Censored at randomisation day + 1 (no tumour assessments while on treatment)	15	12.2	13	9.8	28	10.9
More than 18 weeks from the last assessment between last evaluation without PD and documentation of PD	1	0.8	1	0.8	2	0.8
Total	123	100.0	133	100.0	256	100.0

IA, Investigator assessment; PD, progressive disease; PLD, pegylated liposomal doxorubicin.

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]	[redacted]		[redacted]		[redacted]	
	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[redacted]

The TAG also points out the same issues related to the OS analysis:

"In addition, limited details around censoring within the OS analysis have been presented within the MS. The manufacturer reports that a total of 114 patients were censored in the final OS analysis for platinum sensitive patients (approximately 27%); however, it is unclear for what reasons, and in which arm these patients were when censored." (pg 253)

*"As a result of the high degree of censoring, and the lack of information provided around the reasons for censoring, the TAG notes that censoring within this analysis may be informative. The TAG notes that the presence of informative censoring may reduce the validity of the Kaplan Meier data presented and used within the model. The TAG does note, however, that censoring is ******** (in terms of both number of patients and proportion of patients) between the two arms of the study on aggregate, although it is unclear at what time points censoring occurred." (pg 254)*

Based on this comment, PharmaMar has conducted an additional analysis to demonstrate the breakdown of the OS for the platinum sensitive patients.

[redacted]

[redacted]

	[redacted]				[redacted]	
	[redacted]		[redacted]			
	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[redacted]

[redacted]	[redacted]				[redacted]	
	[redacted]		[redacted]			
	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]	[redacted]				[redacted]	
	[redacted]		[redacted]			
	[redacted]		[redacted]			
	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[redacted]

[redacted]	[redacted]				[redacted]	
	[redacted]		[redacted]			
	[redacted]		[redacted]		[redacted]	
	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[redacted]

[redacted]

Issue 2 The TAG notes that the assessment of a statistical significant difference in PFI was estimated based on the full population (i.e. including platinum resistant/refractory patients) rather than for platinum sensitive patients alone

<table>
<tr><th>Description of problem</th><th>Description of proposed amendment</th><th>Justification for amendment</th></tr>
<tr>
<td>
In the Assessment Report consultation “the TAG notes that the assessment of a statistical significant difference in PFI was estimated based on the full population (i.e. including platinum resistant/refractory patients) rather than for platinum sensitive patients alone.” (pg 253)

On page 18 of the manufacturer’s submission, the statistical difference is reported for both the full population and the platinum sensitive population through the p value.

That is:
<table>
<tr><th></th><th>Trabectedin plus PLDH</th><th>PLDH monotherapy</th><th>P value</th></tr>
<tr><td>Mean PFI (all patients within OVA-301)</td><td>10.6 months</td><td>13.3 months</td><td>P = 0.009</td></tr>
<tr><td>Mean PFI (platinum sensitive patients within OVA-301)</td><td>14.3 months</td><td>19.0 months</td><td>P = 0.002</td></tr>
</table>
</td>
<td>PharmaMar requests that the TAG acknowledge that statistical significant difference in PFI was provided for both the full population and the platinum sensitive patients alone.</td>
<td>To ensure the TAG and AC have all available information to inform decision making.</td>
</tr>
</table>

Issue 3 The TAG requests clarification regarding the derivation of HRQoL

<table>
<tr><th>Description of problem</th><th>Description of proposed amendment</th><th>Justification for amendment</th></tr>
<tr>
<td>
The TAG requested clarification around the derivation of HRQoL from OVA-301:

“Although not reported, the TAG considers it likely that the estimates of HRQoL were derived from the full population included within the OVA-301 trial; i.e. both platinum-sensitive patients and platinum resistant patients.” (pg 255)

The mean utilities used in the stable and progressive states were 0.718 and 0.649, respectively [ScHARR 2010]. These were used following the recommendation from the ERG during the original STA: “the ERG believes that the most sensible health state utility values are those originally presented in the MS, i.e. health state utilities estimated for the combined treatment group across all platinum-sensitive women (Table 28)” [ScHARR 2010].

Table 28: Health state utilities used in the original MS (reproduction of table 36, p 141 in the MS)
<table>
<tr><th></th><th>Mean</th><th>Std Error</th><th>Lower 95% CI for Mean</th><th>Upper 95% CI for Mean</th></tr>
<tr><td>Stable</td><td>0.718</td><td>0.01</td><td>0.699</td><td>0.737</td></tr>
<tr><td>Progressive</td><td>0.649</td><td>0.019</td><td>0.611</td><td>0.686</td></tr>
</table>

As described above these utilities were derived for the combined treatment group across all platinum sensitive women (PFI > 6 months) within the OVA-301 trial. This does not include platinum resistant women.
</td>
<td>PharmaMar requests that the TAG acknowledges the clarification provided on the derivation of HRQoL.</td>
<td>To ensure the TAG and AC have all available information to inform decision making.</td>
</tr>
</table>

Issue 4 The TAG suggests that the discounting methodology overestimates the QALY gain in the manufacturer's model

Description of problem	Description of proposed amendment	Justification for amendment
PharmaMar would like to emphasise that the model structure presented as part of the manufacturer's submitted evidence was based solely on the structure created by a NICE TAG during TA91. Cost-effectiveness estimates from this model structure have been shown to be robust enough to inform NICE recommendations for therapies within relapsed ovarian cancer during TA91. We therefore challenge that by suggesting the submitted model structure and discounting methodology do not provide plausible estimates of cost-effectiveness, there is inconsistency between what is deemed acceptable in TA91 and the review in progress. We acknowledge the criticism by the ERG involved in TA222 which has been noted by the BMJ-TAG: *"Discounting cannot be easily implemented in such a model structure. Ideally a state transition-type Markov trace element should be constructed to facilitate the implementation of discounting." (pg 363)* Although this highlights that it is more challenging to implement discounting in such a model compared to standard cyclic discounting in a Markov trace, we would point out that this critique does not suggest discounting is incorrect or inappropriate when implemented within the model structure. Nor does it suggest that QALY gains are overestimated by such a method, which has been extrapolated by the TAG. Furthermore, we would highlight that we used the "corrected approach" for discounting recommended by the ERG (Sheffield University) during TA222 [Papaioannou et al. 2011]. As a consequence of the TAGs preference to facilitate the implementation of discounting using monthly cycles, the manufacturer's discounting method using the exponential distribution with an instantaneous rate of 3.44% was criticised: *"The TAG economic analysis did not rely upon mean estimates of PFS or OS. Instead, costs and QALYs were estimated monthly for each health state; these costs and QALYs were then discounted depending upon the year in which they fell.* *The TAG considers that, as a result of the discounting methodology used, the manufacturer may have overestimated the QALY gain. This is because, application of discounting to average estimates is unlikely to be as accurate as discounting based on monthly estimates, as the granularity of patient proportions, by health state, over time, is not captured." (pg 363)*	PharmaMar request that the TAG amend the report to state that the discounting method does not overestimate the QALY gain in the manufacturer's model, and that by using the exponential or cyclic methods of discounting similar results are observed.	The interpretation that QALYs gained is overestimated by the manufacturer's model is unfounded. Considering the Excel example (Appendix 1) and updated cost-effectiveness model (Appendix 2) similar results are observed.

We challenge the idea that discounting overestimated the QALY gain, and that the TAG's method is more accurate:
• Firstly, if QALY gain is overestimated then one would equally expect costs to be overestimated as well since the model uses the same method of discounting for both costs and QALYs. However, the TAG suggests that *"The difference in ICER estimated between the TAG model and the manufacturer's model was a result of a difference in estimated incremental QALYs, rather than the difference in incremental costs." (pg 33)* This appears to give an inconsistent message that the discounting methodology only affects the QALYs.
• Secondly, we would argue that the exponential method actually provides (slightly) more accurate results since costs and QALYs over time are discounted continuously as opposed to monthly intervals with half-cycle correction. One would expect the "granularity of patient proportions, by health state, over time" to be more accurate with continuous time than partitioned cycles.
• Thirdly, we provide a simple example in Excel (Appendix 1) to show that discounting costs and QALYs continuously by the exponential method with an instantaneous rate of 3.44% should give no noticeable differences compared to cyclic discounting at 3.5% per annum. Although, the cyclic discounting method appears to produce marginally lower costs and QALYs.
• Finally, to prove the method of discounting in the model and use of mean time in state does not significantly impact results compared to a monthly cyclic Markov model, we provide an updated cost-effectiveness analysis (Appendix 2). The amends to the model include:
  ○ The addition of a 'Monthly Cycles' sheet which calculates discounted costs and QALYs each month over a lifetime horizon (50 years) at a 3.5% rate p.a.
  ○ In particular, for each health state the proportion of patients, costs and QALYs are calculated each month for trabectedin plus PLDH and PLDH monotherapy.
  ○ This analysis results in a £28,402 ICER with PAS which is a small increase compared to the £27,573 ICER with PAS submitted in the manufacturer's submission (shown on the 'Analysis' sheet in Appendix 2). Incremental costs and QALYs slightly reduced when considering cyclic discounting (£11,681 and 0.41) compared to the originally submitted model (£13,397 and 0.49). This aligns with the observations in Appendix 1. As discussed, although the ICER produced by cyclic discounting may be slightly more imprecise, it does not appear to largely affect the final ICER.
  ○ As a result the difference in ICER estimated by the TAG, £35,646 and the manufacturers		

ICER of £27,573 cannot be due to the method of discounting alone as the TAG suggests : *"The ICER estimated by the TAG in this exploratory analysis was £35,646 per additional QALY. This figure compares with the manufacturer estimate of £27,573 for the same comparison. The difference in ICER estimated between the TAG model and the manufacturer's model was a result of a difference in estimated incremental QALYs, rather than the difference in incremental costs. The TAG believes that this may be due to the manufacturer's method of discounting future costs and benefits;" (pg 33)* ○ Indeed, **the most likely reason for the difference is explained under Issue 2** within PharmaMar's comments on the executable model.		

Issue 5 Importance of adjustment of PFS and OS data using PFI as a continuous variable and interpretation of cost-effectiveness results

Description of problem	Description of proposed amendment	Justification for amendment
As discussed in the manufacturer's submission, adjusting PFS and OS data for potential covariates was explored. This analysis found a noticeable difference between treatment arms including a 4.7 month (25%, p=0.002) difference in PFI. As a result PFS and OS data were adjusted by covariates including PFI. A scenario analysis was conducted to identify the impact of not adjusting for PFI and results proved to be very sensitive with significantly worse fitting survival distributions. For the base case, the TAG does not adjust for imbalances observed in the PLDH arm, which is acknowledged as an important driver of results: *"However, the TAG considers that the difference in the ICER between the TAG's and manufacturer's base case analyses is predominantly a consequence of the use of adjusted clinical effectiveness data; adjusted for baseline characteristics such as PFI (as a continuous variable). The TAG notes that adjustment of clinical effectiveness data for key prognostic factors such as PFI is likely to result in more accurate estimates of PFS and OS." (pg 363)*	PharmaMar request that the TAG clarify which cost-effectiveness estimates should be used to inform cost-effectiveness decisions for trabectedin plus PLDH vs. PLDH monotherapy. Although the non adjusted ICERs provide a comparable dataset, following the discovery	It is unclear whether adjusted or unadjusted cost-effectiveness for trabectedin plus PLDH vs. PLDH monotherapy should be used to inform cost-effectiveness decisions.

We are pleased that the TAG have acknowledged this weakness throughout the Assessment Report, and that the scenario analysis which does adjust for these imbalances is considered to be the most accurate cost-effectiveness estimate for trabectedin plus PLDH vs. PLDH monotherapy: *"The TAG considers that the ICER estimated in the TAG's scenario analysis is likely to be the most accurate reflection of the cost-effectiveness of trabectedin plus PLDH versus PLDH" (pg 363)* However, we feel that there is some confusion as to which ICER the TAG believe is more appropriate to evaluate the cost-effectiveness of trabectedin plus PLDH vs. PLDH and other non platinum agents. The TAG state that: *"However, as efficacy data used in the TAG's base case model was unadjusted (to provide a consistent dataset), the TAG notes that the head-to-head ICER generated from using adjusted efficacy data is not comparable with ICERs estimated for other treatments in the TAG's base case analyses." (pg 30)* *"The TAG notes that, given the data available, it is not possible to adjust survival for the remaining treatments in the network in a similar way, and therefore the results of the scenario analysis should be considered independently." (pg 35)* These statements seem at odds with the interpretation that the scenario analysis offers the most accurate reflection of cost-effectiveness for trabectedin plus PLDH vs. PLDH monotherapy. We acknowledge that whilst the probabilistic ICER for trabectedin vs. PLDH without adjusting for covariates (£81,353) is considered more consistent and comparable with other non platinums, it is significantly inaccurate compared to the ICER calculated in the scenario analysis (£35,646). We would therefore recommend to the ERG a two stage approach: 1. Firstly, a consistent approach is applied to all non platinums by not adjusting for covariates. 2. Secondly, following the discovery that PLDH is indeed the most cost-effective of the non platinums, the scenario analysis should be considered to provide the most accurate reflection of the cost-effectiveness of trabectedin plus PLDH vs. PLDH monotherapy. 3. Hence, recommendations may be made with regards to the cost-effectiveness of non platinums when not adjusting for covariates, and the cost-effectiveness of non platinums when correctly adjusting for covariates. In this regard, the ICER of trabectedin plus PLDH vs. PLDH monotherapy, when covariates have been adjusted for, is the ICER for which decision makers should be basing cost-effectiveness decisions on.	that PLDH is the most cost-effective treatment under this assumption, the adjusted ICER of trabectedin plus PLDH vs. PLDH monotherapy is more appropriate and should be used to inform cost-effectiveness decisions.	

Issue 6 The TAG requests justification of curve fitting separately by treatment arm

Description of problem	Description of proposed amendment	Justification for amendment
The TAG indicated that no justification was provided for fitting parametric distributions to the Kaplan Meier data by treatment arm compared to selecting treatment as a covariate: *"These curves were fitted separately by treatment arm; i.e. treatment was not included as a covariate; the manufacturer did not provide a rationale within the submission for this methodology."* (pg 27) PharmaMar acknowledge that the rationale for this selection could have been clearer. It is important to note that all statistical analyses with regards to PFS and OS were conducted in line with the methodology recommended by the ERG of TA222 (Sheffield University): [Papaioannou et al. 2011] "Distributions were fitted to each treatment separately and adjusting for potential covariates" [Papaioannou et al. 2011] The ERG during TA222 agrees with the TAG in that the proportional hazards assumption "may not be the case in real life" [Papaioannou et al. 2011]. In addition, when considering the cumulative log hazards associated with Kaplan Meier PFS and OS (Appendix 10 of the TAG Assessment Report), it is clear that hazards are not proportional and thus fitting treatment as a covariate to the parametric curves would inappropriately assume proportional hazards. This limitation is discussed in great length within the TAG Assessment Report. Additionally, the DSU TSD provides further justification as to why curves should be fitted by treatment arm and not using treatment as covariate (assuming proportional hazards): [NICE DSU 2013] • Inspection of the cumulative log hazards suggests the proportional hazards assumption does not hold. • Proportional hazards cannot be assumed for the log-logistic parametric distribution, which has been shown to be a significantly better fit to the OS Kaplan Meier data for trabectedin plus PLDH and PLDH monotherapy based on AIC/BIC (see Appendix 7 of manufacturer's submitted evidence). Hence less accurate models for OS would have to be selected. • "If the proportional hazards assumption does not seem appropriate it is likely to be most sensible to fit separate parametric models of the same type" [NICE DSU 2013]. This is exactly what we did for	PharmaMar request that the TAG acknowledge that following consultation, the manufacturer has provided a clear rationale for choosing to fit curves separately by treatment arm, which mitigates the bias of assuming proportional hazards and provides more reflective estimates of PFS and OS.	To ensure the TAG and AC have all available information to inform decision making.

both PFS and OS. We are pleased that the TAG have acknowledged that the manufacturer's statistical analysis used during the TAG's scenario analysis provides the most accurate cost-effectiveness estimate for trabectedin plus PLDH vs. PLDH monotherapy: *"The TAG considers that the ICER estimated in the TAG's scenario analysis is likely to be the most accurate reflection of the cost-effectiveness of trabectedin plus PLDH versus PLDH" (pg 363)*		

References

Data on file. EMEA response. Confidential data.

School of Health and Related Research (ScHARR). Trabectedin for the treatment of relapsed ovarian cancer: A Single Technology Appraisal. April 2010. Available from: http://www.hta.ac.uk/erg/reports/2215.pdf

Papaioannou D, Rafia R, Stevens J et al. Trabectedin for the treatment of relapsed ovarian cancer: A single Technology Appraisal. Health Technology Appraisal 2011; 15: 69-75.

NICE Decision Support Unit. Survival analysis for economic evaluations alongside clinical trials – extrapolation with patient-level data. March 2013. Available from: http://www.nicedsu.org.uk/NICE%20DSU%20TSD%20Survival%20analysis.updated%20March%202013.pdf