Comments on The clinical and cost effectiveness of topotecan for small cell lung cancer: a systematic review and economic evaluation

We welcome the opportunity to comment on the assessment report produced by the Southampton Technology Assessment Group (SHTAG) for the Health Technology Assessment of topotecan for the treatment of Small Cell Lung Cancer.

This document is divided in two sections:

- Section A contains specific comments on the technical content of the clinical and cost effectiveness section of the report.

- Section B attempts to address the concerns raised by the SHTAG when appraising GSK's cost effectiveness analysis and modelling approach

SECTION A GSK comments on Technology Assessment Report (TAR)

1 Clinical effectiveness section

- The conclusions of the executive summary reads ***"In summary, the clinical evidence indicates that topotecan is better than BSC alone in terms of improved survival, is as effective as CAV, and less favourable than IV amrubicin in terms of response. Oral topotecan and IV topotecan were shown to be similar in efficacy. It remains uncertain whether topotecan is more or less toxic than comparator interventions."*** However, we feel that the wording of the statement on IV amrubicin (underlined above) needs some consideration given the limited internal and external validity of the clinical trial supporting it (Inoue et al, 2008). Please see below for a more detailed discussion of the main limitations of the Inoue et al study.

- Among the main limitations of this study there are two which are believed to have the greatest impact on its external validity and need to be taken into account when extrapolating the study results to a UK population:

- Small sample size: Only 60 Japanese patients were enrolled in this Phase II study
- Participants randomized to the IV topotecan arm received a lower dose of IV topotecan (1.0 mg/m^2) compared to the licensed UK dose of 1.5 mg/m^2. Thus the dose used was two-thirds that recommended in the UK.
- Due to well documented inter ethnic differences of dose response between Japanese and Caucasian populations, the outcomes of this study are not directly transferable to the UK population. The International Conference on Harmonization (ICH) published the "Guidance on Ethnic Factors in the Acceptability of Foreign Clinical Data" in 1998. In this report the ICH stated that one of the most important determinants in extrapolating clinical data from one region to another involves ethnic differences in the pharmacokinetics of the drug.

2 Cost effectiveness section

2.1 General overview

The health economic model submitted by GSK is based on individual patient level data from the O'Brien et al. trial whilst the SHTAG model is essentially an economic model based on aggregate data, also derived from data from the same O'Brien et al. trial. All else being equal, it can be argued that an approach using individual level data is likely to be a more accurate representation of the subjacent clinical data (with respect to survival, quality of life and costs) and therefore it constitutes a more robust modelling approach which yields a more representative ICER.

Please note that initial concerns raised by the SHTAG on the modelling approach have now been addressed in section B of this document.

2.2 Survival modelling

With regard to the survival curves (p.65 and following of the SHTAG report), it is not clear to us what benefit is gained from modelling survival using the parametric survival curve as opposed to the Kaplan-Meier method (the latter most closely reflecting the individual level approach reported in the manufacturer submission), especially in the light of the additional sensitivity analysis reported in table 1 of section B. While the parametric approach does yield more favourable estimates for oral topotecan this would appear to

be driven by an overestimate of the survival curve in the period 50-100 weeks. Hence, it can be argued that an approach using individual level data is likely to be a more accurate representation of the subjacent clinical data (with respect to survival, quality of life and costs) and therefore should be preferred.

2.3 HRQoL

In terms of measuring HRQoL both the SHTAG and manufacturer approaches essentially rely on predicting EQ-5D scores beyond the 12th cycle. However, the SHTAG analysis (p.70 and following of the assessment report) was based on aggregate data which should not be considered as robust as an analysis using individual level data; especially in the light of the additional sensitivity analysis reported in table 1 of section B.

2.4 Cost calculation

The cost analysis in the SHTAG report is similar to that in the manufacturer submission and yields very similar results for most cost components. However, there are some discrepancies with respect to administration costs and monitoring costs for oral topotecan (p.74 and following) and some clarification of the rationale behind the cost analysis would be very welcome.

There are two specific concerns:

1. Some of the costs attributed to drug administration and monitoring in the SHTAG report are in reality attributable to the disease not the drug and would equally apply to patients receiving Best Supportive Care only and so should not be included. More specifically, relapsed SCLC patients receiving best supportive care only would also need to be medically evaluated and receive routine test (e.g. FBC, chest X-ray, U&E) as part of the regular monitoring of their disease

2. In terms of the size of the monitoring costs, there seems to be an error in the calculations presented in the SHTAG report. The assessment report assumed that patients receiving active treatment would have a CT scan every two cycles. (Page 74 -75). However, when calculating the monitoring cost for the complete treatment duration of four courses/cycles of chemotherapy, it appears that cost of a CT scan is calculated as it was given every cycle not every two cycles (see Table 28 from assessment report and paragraph below – page 75)

SECTION B Response to “Summary of general concerns” (page 62)

“It is unclear whether the disutility that would be expected from experiencing an adverse event in the topotecan group has been adequately represented due to the large amount of missing EQ-5D data and three week intervals between collections of EQ-5D data. This may be further biased due to healthier participants being more able and willing to fill in EQ-5D questionnaires than those who are experiencing an adverse event. If this is correct, then utility and therefore gain in QoL compared to BSC is likely to be an overestimation for the topotecan group.”

We agree with these concerns. It is noted that they apply equally to the SHTAG analysis which uses the same data

“No modelling beyond the length of the trial was undertaken. A small but potentially significant number of participants were still alive at the end of the trial. However, it is not entirely clear how many participants in the trial were still alive, as the MS and Kaplan-Meier plot from the O’Brien and colleagues RCT57 seem to give conflicting reports. It is assumed here that the MS is correct as the participant level data is given in the model. Therefore, just over 4% of each arm of the trial were still alive at the end of the study and there is a possibility this could have underestimated the survival benefit for either group.”

A further sensitivity analysis is reported below that assesses this issue. Time to death was varied among the 6 patients alive at the point of their final follow-up visit. The baseline estimate assumes that all 6 patients died the day after this point. The sensitivity of the results to this assumption was investigated, considering two alternative scenarios. In the first scenario it was assumed that the 3 censored BSC alone patients died the day after the final follow-up visit and that the 3 censored OT+BSC patients survived for exactly one year beyond the final follow-up visit. In the second scenario these changes were reversed, assuming that the 3 censored BSC alone patients survived for exactly one year beyond the final follow-up visit and that the 3 censored OT+BSC patients died the day after the final follow-up visit. These changes affected the QALY estimates only, not the costs; the changes affect the period after treatment and the only relevant costs are those pertaining to additional non-progressive disease survival. Since all 6 patients were in progressive disease before they were censored, these costs are assumed to be unchanged. The results of the sensitivity analysis are summarized in Table 1. The values are within the range of the original sensitivity analysis (e.g., £22,512-£40,253).

Table 1. Additional sensitivity analysis of the cost-effectiveness of oral topotecan + BSC versus BSC alone (all patients).

Scenario	Incremental cost per QALY gained, UK£
All patients	
Baseline estimate*	**26,833**
Survival of censored patients	
Censored BSC alone patients assumed to die the day after the final follow-up visit**;** censored OT+BSC patients assumed to survive for one year beyond final follow-up visit	**23,474**
Censored BSC alone patients assumed to survive for one year beyond the final follow-up visit; censored OT+BSC patients assumed to die the day after the final follow-up visit	**29,206**

OT=oral topotecan, BSC=best supportive care, QALY=quality adjusted life year.

* 6 patients alive at the point of the final follow-up visit were all assumed to die the day after that point.

"The use of the mean observed EQ-5D scores from both arms of the trial to take account of the missing EQ-5D data raises a number of problems. Utility in both groups of participants in the trial is unlikely to be the same throughout the cycles. The utility for topotecan participants early in the treatment cycles is likely to have been underestimated, as this is when the majority of BSC participants were progressing towards death. In the latter half of the treatment cycles the mean of the observed EQ-5D scores appear to have been overestimated, due to the small number of observations and as the proportion of healthier participants increases. It is not clear what effect this will have had on the model results."

The assumption described above was applied in a tiny proportion of cases and any changes will not affect on the results. To clarify, across the 141 patients in the study there were 1,548 21-day survival periods. Individual EQ-5D data were available for 600 periods (39%). Of the 948 missing values there were gaps where EQ-5D data were not collected at every cycle up to cycle 12 for only 4 periods (i.e., <1%). It is only in these cases (i.e., 4 out of 948, <1%) that the mean values across both groups were used.

"The assumptions over the costs in the model appear reasonable. Given that costs for the BSC arm of the trial were not recorded and that this component is common to both arms the conservative assumption may be justified. However, a small percentage of palliative care costs

are likely to have occurred in different periods for the topotecan and BSC and BSC alone groups and discounting could have been applied here. "

We agree with this point, though as evidenced by the SHTAG report the impact of applying discounting is negligible.

"The description of how utilities were used in the model and the methods by which EQ-5D values were imputed to allow for missing data were not entirely clear in the MS."

Below is a more detailed explanation, plus some further sensitivity analyses.

Individual HRQOL data were recorded for every patient in the trial at the start of each cycle up to and including cycle 12. Patients recorded their health status using the EQ-5D. To calculate quality adjusted survival for each patient individual survival in each 21-day period was multiplied by their EQ-5D score pertaining to that period. Across the 141 patients in the study there were 1,548 21-day survival periods. Individual data were available for 600 periods (39%). There were 948 missing values because: there were gaps where EQ-5D data were not collected at every cycle up to cycle 12 (=4 periods [<1%]); EQ-5D data were not collected beyond cycle 12 (=46 periods [3%]); and the recording of EQ-5D data during the first 12 periods ended prematurely, usually as the patient progressed towards death (=898 periods [58%]).

Missing values were imputed using data from the trial. For gaps up to cycle 12 (=4 periods [<1%]) the combined mean EQ-5D score at each cycle was inserted (Table 2). For all other missing values the mean EQ-5D scores in the BSC group after disease progression for the 5 21-day survival periods immediately prior to death was calculated (Table 3).

For oral topotecan+BSC patients a last observation carried forward (LOCF) approach was used until the period they entered progressive disease (PD). Data in Table 3 were applied backwards from the period in which the patient died. If the patient survived more than 5 periods in PD, the figures for the 4th from last period before death were applied backwards until the start of PD.

For BSC alone patients a LOCF approach method was used until 5 periods from death and then data in Table 3 were applied.

Table 2. Combined mean EQ5D scores

Cycle/period	Baseline	1	2	3	4	5	6	7	8	9	10	11	12
EQ5D score	0.657	0.585	0.652	0.646	0.633	0.684	0.690	0.796	0.733	0.689	0.568	0.620	0.620
n	134	112	92	75	55	35	18	7	6	4	2	1	1

Table 3. EQ-5D scores after disease progression (based on estimates from BSC alone patients only).

Period	4th from last period before death	3rd from last period before death	2nd from last period before death	Last period before death	Period in which patient died
EQ5D score	0.658	0.608	0.544	0.250	0.250
n	12	20	29	38	38

To calculate these values we used EQ-5D data for BSC alone patients whose last recorded EQ5D scores was within 6 weeks of death. We grouped the data into 21-day survival periods. We then computed the mean EQ-5D score in each period, combining the data for the period in which the patient died and 1 period before death due to the high number of missing values for the period in which the patient died. We only used data for up to 4 periods before death because the number of observations was very small at 5 or more periods before death (i.e., very few BSC alone patients survived more than 5 periods).

BSC EQ-5D data was used after disease progression for both oral topotecan and BSC patients for the following reasons. First, on balance it appears that oral topotecan (OT) is associated with an improvement in HRQOL over BSC alone during the period in which EQ-5D scores were collected. If OT+BSC EQ-5D data for the OT+BSC group and the BSC alone EQ-5D data after disease progression for the BSC alone group had been used then the higher EQ-5D scores achieved in the OT+BSC during the OT intervention period would have been transmitted into the period of disease progression. We believe this may potentially have introduced a bias in favour of OT+BSC and therefore the approach taken in our submission avoided this potential bias.

Second, little data on EQ-5D scores among OT+BSC patients in progressive disease leading up to death were available. This is because in the trial EQ-5D data were only collected for up to 12 cycles and many OT+BSC patients did not enter progressive disease until after this point. In addition, even fewer of these patients died during the period in which EQ-5D data were collected. The upshot of this is that while comparable estimates to those in Table 3 for OT+BSC patients in progressive disease can be computed, the sample size is small and it is not possible to produce as along a time series back from the period in which death occurred.

Taken together, the above points led us to believe that using OT+BSC EQ-5D data for OT+BSC patients would not be appropriate. Nonetheless, the figures generated using this approach are in Table 4. These are analogous to the figures in Table 3, but estimated on OT+BSC patients in progressive disease. Note the small sample size. Note as well that for the limited periods that are covered the estimates are higher than those in Table 3. Both of these points are consistent with the arguments above.

While, in our view, is not possible to compute satisfactory EQ-5D scores for OT+BSC patients in progressive disease, EQ-5D scores have been computed where they are available for OT+BSC patients immediately prior to death. These are shown in Table 5. Note that the patients from which these data were taken were not necessarily in progressive disease. Therefore, these figures might overestimate the effect of being in progressive disease among OT+BSC patients.

We present cost-effectiveness estimates using the data in Tables 4 and 5 in Table 6. In both cases these values result in more favourable (i.e., lower) cost-effectiveness ratios for OT+BSC.

Table 4. EQ-5D scores in progressive disease (based on estimates from OT+BSC patients only*).

Period	4th from last period before death	3rd from last period before death	2nd from last period before death	Last period before death	Period in which patient died
EQ5D score	.	.	0.689	0.404	0.404
n	0	0	1	14	14

* Note that these figures only include patients in progressive disease.

Table 5. EQ-5D scores leading up to death (based on estimates from OT+BSC patients only*).

Period	4th from last period before death	3rd from last period before death	2nd from last period before death	Last period before death	Period in which patient died
EQ5D score	0.6407	0.5715	0.5885	0.4745	0.4745
n	9	13	15	20	20

* Note that these figures include patients not in progressive disease.

Table 6. Sensitivity analysis of the cost-effectiveness of OT+BSC versus BSC alone.

Scenario	Incremental cost per QALY gained, UK£
Baseline estimate	26,833
HRQOL	
Estimates of EQ-5D scores leading in progressive disease for OT+BSC group based on estimates from OT+BSC patients only,§ and estimates for BSC alone group based on estimates from BSC alone group only	24,718
Estimates of EQ-5D scores leading up to death for OT+BSC group based on estimates from OT+BSC patients only,¶ and estimates for BSC alone group based on estimates from BSC alone group only	25,748

OT=oral topotecan, **BSC**=best supportive care, **QALY**=quality adjusted life year.

§ The EQ-5D values for OT+BSC patients are based on figures only including patients in progressive disease (i.e., using the data in Table 4)

¶ The EQ-5D values for OT+BSC patients are based on figures including patients not in progressive disease (i.e., using the data in Table 5)

In conclusion, we believe that a modelling analysis using individual level data (GSK approach) is likely to be a more accurate representation of the subjacent clinical data - with respect to survival, quality of life and costs - when compared to an economic model based on aggregate data (SHTAG approach) and therefore it would constitute a more robust modelling approach (all else being equal) which should be preferred.