AChEIs and memantine for Alzheimer's Disease

PenTAG responses to Consultee comments

17th August 2010

Given the large volume of comments from the manufacturers and other consultees, we do not intend to give a comprehensive point-by-point response to all comments. Instead we have chosen to address:

- Those recurring comments raised by a number of consultees where we think our clarification would help the Appraisal Committee's deliberations.
- Those assertions that our cost-effectiveness modelling is flawed or contains important technical errors.

We stress that when we adjust our model in response to the manufacturers' criticisms in point 1 below, the cost-effectiveness of all drugs improves substantially.

Economic evaluation

1. Consistency of equations for MMSE as a function of time to institutionalisation and treatment effect in PenTAG's model and discrepancies in results of statistical analysis of time to institutionalisation performed by PenTAG and by Wolstenholme et al. (2002)

Criticism

Eisai/Pfizer state on p14: "*Finally, we note that there is a logical inconsistency in the equations. PenTAG developed equations which allow for the prediction of MMSE by time to institutionalisation. For mild to moderate disease, these equations indicate that a 1 year difference in time to institutionalisation is associated with a 4.17 difference in MMSE. The same equation (MMSE = 8.34 + 4.17t) can be used to solve for time to institutionalisation for a given MMSE score. A 1 point difference in MMSE would result in an almost 3 month delay in institutional care. This is inconsistent with the time to institutionalisation equation, which predicts much smaller delays. Not only is this inconsistent, but it leads to an incoherent model. A 1 point change in MMSE will lead to roughly a 10 day delay in institutional care according to the PenTAG equation. This 10 day delay is then used to calculate differences in MMSE scores in the pre-institutional care state. The 10 day delay, however, means that the original 1 point difference in MMSE is transformed to a difference equal to 4.17*10/365, or 0.11.*"

In a related point Lundbeck state on p18 and Eisai/Pfizer state on p12 that Wolstenholme et al (2002) found that MMSE and Barthel were significant predictors of time to institutionalisation, whereas we did not.

Eisai/Pfizer state on p13 that: "*Although MMSE and ADL scores were available for individuals over the course of the study, PenTAG opted only to use baseline MMSE and ADL as potential predictors. A more appropriate approach would have evaluated how changes in MMSE/ADL, and 'current' MMSE/ADL scores influence the risk of institutionalisation*".

Response

We thank the consultees for these comments, and we admit that we have incorrectly modelled the treatment effect for time to institutionalisation and overall survival.

When we corrected our model, the cost-effectiveness of all drugs against BSC improves substantially (see separate updated ICER tables). We stand by our modelling of time to institutionalisation and overall survival for BSC, and indeed this is not contested by the consultees. The first paragraph, in Section 7.3.4. is now incorrect. This paragraph should read;

"The model starts when treatment begins for the treated cohorts (point A in Figure 60). For patients in the BSC arm only, the probability distributions for the time to institutionalization and time to death are predicted using mean baseline characteristics of the cohort, these being mean age at baseline, mean MMSE at baseline and mean Barthel at baseline."

Following recalculation the text should continue as follows to the end of point 1.:

"After the initial treatment period (Figure 60 point B), any treatment effects, in terms of improved MMSE and Barthel, are assumed to have occurred. The treatment effects were translated into improved time to institutionalisation and overall survival in the following manner. First, a linear mixed effects model (from the "nlme" "R" package) was fitted with time to end of pre-institutionalization (or overall survival) as the response variable, and MMSE and Barthel as explanatory variables, and patient as a random effect. For each patient there were typically several observations. Variations in the intercept and slopes of the effects of MMSE and Barthel across patients were modelled as random variables, as normal distributions. In addition, a covariate, MMSE at the start of the study was included in the model. The following equations were obtained;

Time to institutionalisation = -1.086 + 0.0640(MMSE) + 0.2001(Barthel) + 0.0023(Baseline MMSE)(MMSE) - 0.0072(Baseline MMSE)(Barthel)

Time to death = 4.46593 - 0.0843(MMSE) - 0.2874(Barthel) - 0.0025(Baseline MMSE)(MMSE) + 0.0109(Baseline MMSE)(Barthel)

For mild / moderate patients, with a mean baseline MMSE of 17, this gives;

Time to institutionalisation = -1.086 + 0.1032(MMSE) + 0.0781(Barthel)

Time to death = 4.46593 + 0.1270(MMSE) + 0.1021(Barthel)

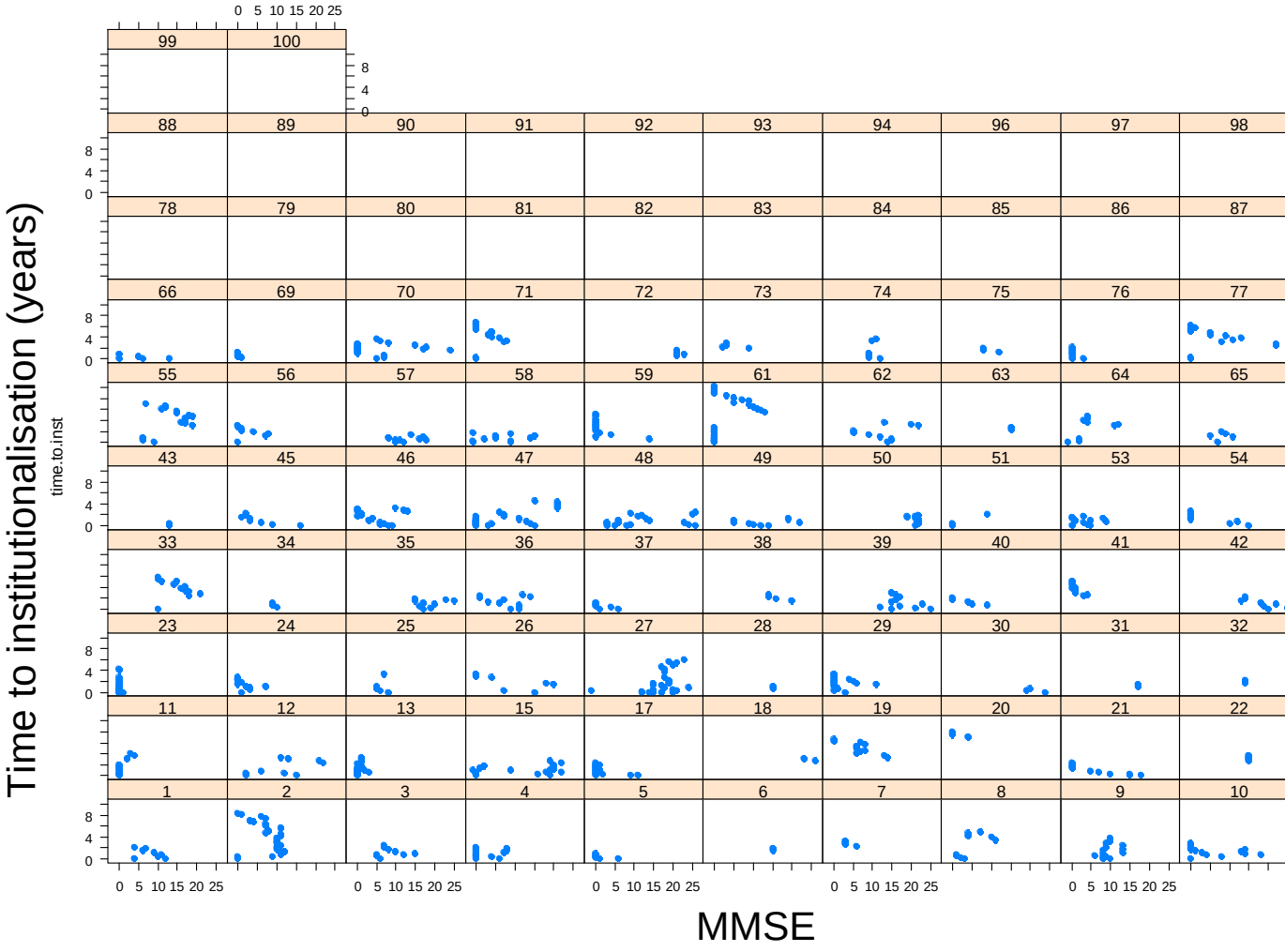
And for moderate / severe patients, with a mean baseline MMSE of 11.7;

Time to institutionalisation = -1.086 + 0.0910(MMSE) + 0.1159(Barthel)

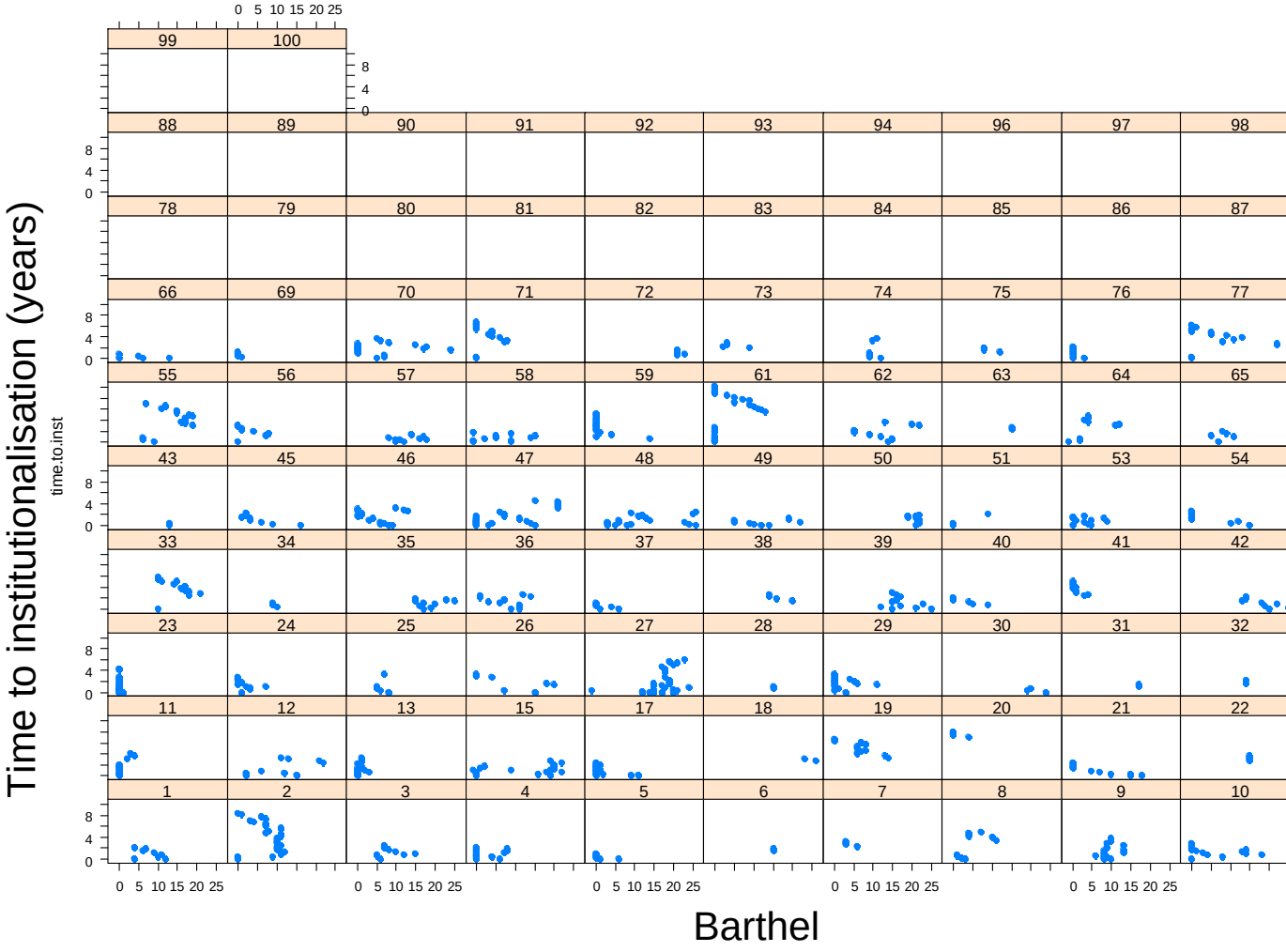
Time to death = 4.46593 + 0.1138(MMSE) + 0.1595(Barthel)

In addition, the uncertainty in these equations was modelled for the PSA by recording the Cholesky matrices from the statistical analyses.

The impact of (time-varying) MMSE on time to institutionalisation can be seen from Figure 1 below, with one panel per patient.



Similarly, the impact of (time-varying) Barthel on time to institutionalisation can be seen from the Figure 2 below.



Similar graphs are produced for the time to death.

From the equations above, the mean increase in the time (years) to institutionalisation for a given drug is calculated for mild to moderate patients as 0.1032(ΔMMSE) + 0.0781(ΔBarthel), and for moderate to severe patients as 0.0910(ΔMMSE) + 0.1159(ΔBarthel), where ΔMMSE and ΔBarthel are the treatment effects on the MMSE and Barthel scales.

Also, the mean increase in the time to death for a given drug is calculated for mild to moderate patients as 0.1270(ΔMMSE) + 0.1021(ΔBarthel), and for moderate to severe patients as 0.1138(ΔMMSE) + 0.1595(ΔBarthel). Note that we do not assume that drugs affect overall survival in the base case. Instead, this is addressed in a sensitivity analysis.

Next, the time to institutionalisation and the time to death for each drug were assumed to follow exponential distributions, as for BSC. The mean time to institutionalisation for a given drug is calculated as the mean time to institutionalisation for BSC plus the mean increase in the time to institutionalisation for the drug. Similarly, for the sensitivity analysis, the mean time to death for a given drug is calculated as the mean time to death for BSC plus the mean increase in the time to death for the drug. The treatment effects in terms of delaying the time to institutionalisation (base case) Figure 3 and overall survival (sensitivity analysis only) Figure 4 are shown for all drugs for mild/moderate disease in the graphs below (for patients starting with mean age of 77), where BSC is shown by the lower lines, and the lines for the four treatments are virtually indistinguishable. Clearly, the impact of the drugs on time to institutionalisation is still small, e.g. delayed by a mean of nearly 2 months for donepezil, and also small on overall survival, e.g. delayed by a mean of 2.3 months for donepezil. However, the delays are substantially greater than in our original analysis.

Figure 3

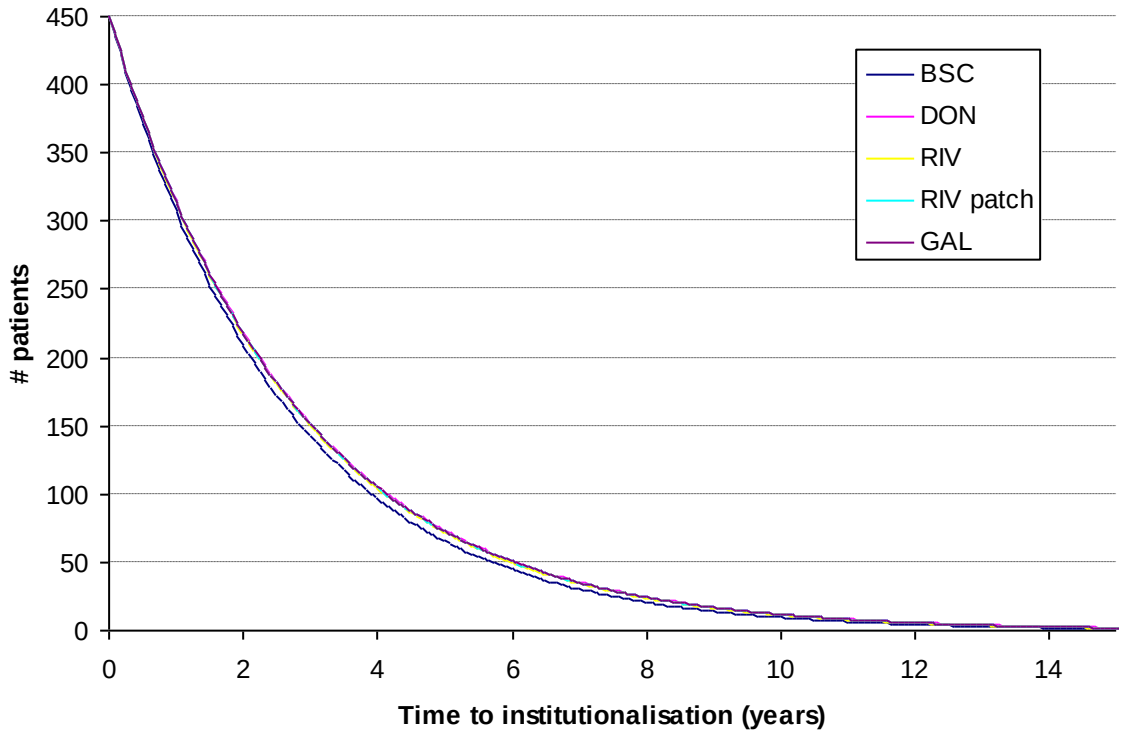
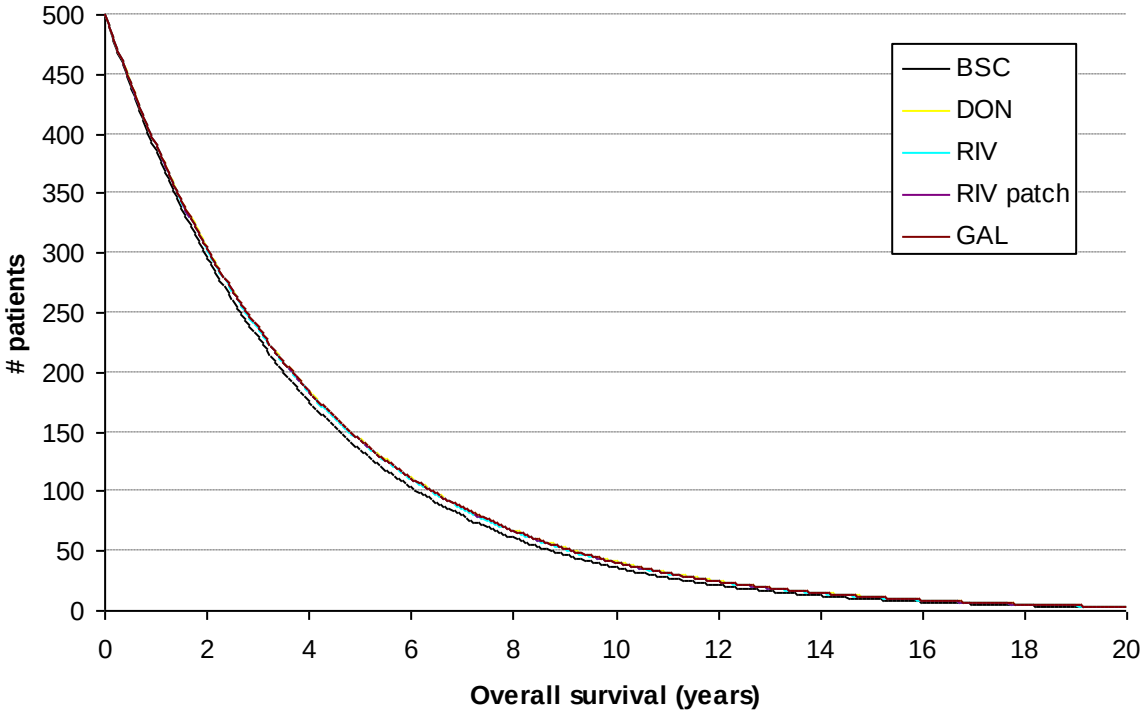


Figure 4

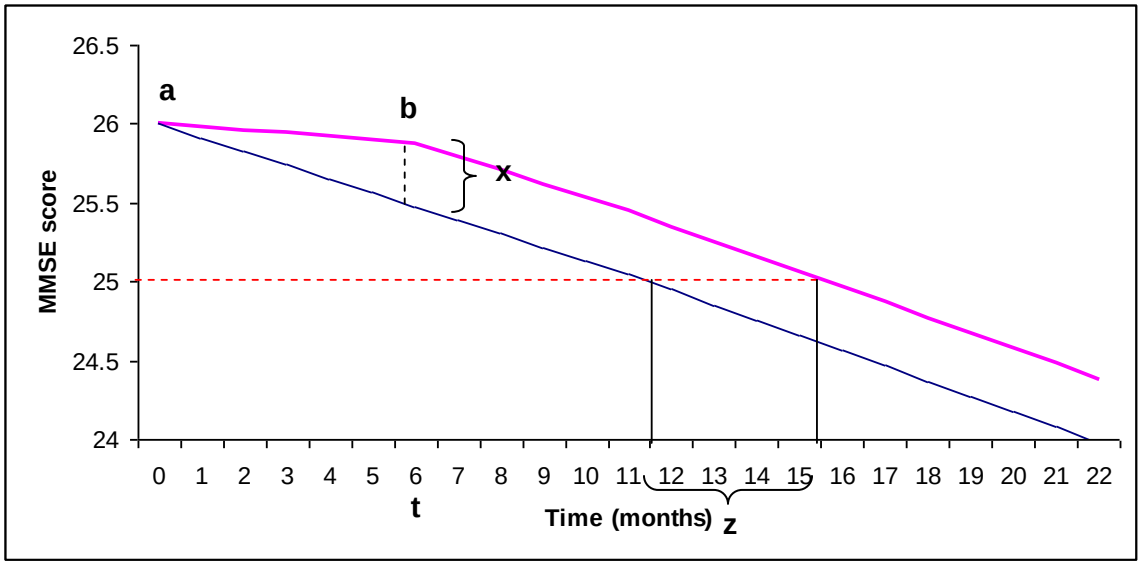


Whilst we admit that we originally modelled the treatment effect incorrectly, it does appear rather paradoxical that the treatment effect we calculated originally was so small. Instead, one might imagine that the treatment effect as calculated originally (where the treatment effects on MMSE and Barthel were assumed to increase the initial MMSE and initial Barthel, and then these revised starting MMSE and Barthel values fed in to the equation in Section 7.3.8 in our report to predict the time to institutionalisation) would be similar to our revised calculation. We believe that the paradox is explained by the fact that the intercepts and slopes in the relationships for MMSE and Barthel versus time to institutionalisation differ substantially between patients, as shown in the panel graphs above.

Implicit in our revised method of modelling the treatment effect on time to institutionalisation and time to death is an important assumption. As shown in the diagram below Figure 5 (Figure 67 Appendix 16 to our report), a typical trajectory for BSC of MMSE over time is shown by the straight lower line, and for a drug, by the upper kinked line. Distance x represents the treatment effect on the MMSE scale (reported in the RCTs), and the time difference z represents the treatment effect in terms of delay in time to institutionalisation or death, as calculated in our revised methodology. The critical assumption is that the slope of the upper line from point b onwards equals the slope of the BSC line, i.e. after six months (the time of follow-up in the RCTs), at which time, some, but not all of the patients have stopped drug treatment, the rate of decline of MMSE is equal for patients on BSC and on the drugs. It could be argued that there could be a "bounce back" effect for times after point b, that is, when patients stop drug treatment, their MMSE declines at a greater rate than patients on BSC. If this were the

case, the cost-effectiveness of all drugs against BSC would be worse than the revised estimates. We are not aware of any data to help further investigate this assumption, i.e. to quantify the difference in MMSE at later times.

Figure 5



Comparison with Wolstenholme et al (2002)

We acknowledge that we should have commented in our report on the apparent discrepancy between our analysis of the Oxfordshire data and that of Wolstenholme et al (2002). The consultees have made a helpful observation, and this has prompted us to correct our method of modelling the treatment effect on time to institutionalisation and overall survival, as explained above. However, we do have some further comments on the comparison of our analysis and that of Wolstenholme et al (2002).

On p19, Lundbeck give a table which shows "predictive coefficients" for Age, MMSE and Barthel from Wolstenholme et al (2002) and PenTAG. They state that the coefficients from Wolstenholme et al (2002) are very different to the values supplied by us. However, we believe that the definitions of the coefficients from Wolstenholme et al (2002) are very different to those from us, as follows.

Wolstenholme et al (2002) used a Cox proportional hazards model to investigate the effects of the covariates on the time to institutionalisation. Table 4 in Wolstenholme et al (2002) gives the hazards for each of the covariates. These are then quoted by Lundbeck in the table on p19 of their report. Importantly, it appears that Wolstenholme et al (2002) have used multiple observations for each patient, that is, they have performed a longitudinal analysis. Whilst this is not clearly stated in the paper, this appears to be the case because in the footnote to Table 4, it states that "no. of subjects = 100, no. of observations = 710".

Now, we turn to our original statistical analysis for the time to institutionalisation. Instead of performing a longitudinal analysis, as appears to be the case in Wolstenholme et al (2002), we modelled time to institutionalisation for BSC as a function of the initial MMSE, initial Barthel and initial age, with one observation per patient. The results of our

analysis are expressed by the equation in Section 7.3.8 of our report. From this equation, we can see that the coefficients are not hazards (as they are in Table 4 of Wolstenholme et al (2002)). As an example to see that the coefficients of Wolstenholme et al (2002) and of PenTAG are different measures, observe that both analyses predict that increasing age decreases the expected time to institutionalisation, however, the coefficient for age is positive in Wolstenholme et al (2002), but negative from PenTAG.

As stated above, the precise statistical model used by Wolstenholme et al (2002) is not clear, although we suspect they performed a longitudinal analysis. As suggested by Lundbeck, we had indeed emailed the person who conducted the statistical analysis for Wolstenholme et al (2002) for clarification of their analysis. However, we received no reply.

N.B. Section 7.3.8. "Health state occupancy", including Figures 64 and 65, is still correct, except we stress that any mention of "MMSE", "Barthel-ADL" is short-hand for "initial MMSE", "initial Barthel-ADL", where "initial" indicates the values at the start of the study. Also, our derivation of a relationship between MMSE and time to institutionalisation in Section 7.3.9.1 'Multiple-state health utility: by cognition, dependency and residential status' is still valid.

2. Fit to time to institutionalisation and overall survival and MMSE over time to institutionalisation in PenTAG model

Criticism

Eisai/Pfizer state on p2 "*Moreover, the equation"* (for time to institutionalisation) *"appears to be mis-specified and shows an extremely poor fit with the observed data from the study. As a result the predictions from the equation are unrealistic.*". This point is repeated on p12 and in Appendix C.

Eisai/Pfizer state on p28: "*Figure 66 shows that the linear fits produced by PenTAG bear no resemblance to the observed data*."

Response

There has been a simple misunderstanding. In Figures 64 and 65, the upper dotted lines represent the upper 95% confidence intervals of the Kaplan-Meier curves and the lower dotted lines represent the lower 95% confidence intervals of the Kaplan-Meier curves. These dotted lines therefore are not to be compared with the fits shown by the continuous smooth curves. However, the Kaplan-Meier best estimates can be compared with the intermediate smooth curves.

Next, we defend our fit of MMSE as a function of time to institutionalisation, as shown in Figure 66 of our report. There is certainly a good deal of variation between patients, as shown by large differences in the lines across patients. However, for cost-effectiveness analysis, we are concerning with the <u>mean</u> relation between MMSE and time to institutionalisation across all patients. This is calculated correctly in our statistical analysis by assuming random effects for the slopes and intercepts. Note that the large spread across patients is reflected in the uncertainty in the line of best fit in our PSA analysis. As an analogy, we assume a single mean utility value for all patients in a certain health state. However, of course there will often be wide variation in utilities across all patients.

3. Use of exponential distributions for time to institutionalisation and overall survival in PenTAG model

Criticism

Lundbeck state on p7, point 1 that the "*assumption that hazard rates would not accelerate over time has no clinical validity*". This point is repeated on p17. Also Eisai/Pfizer state on p13: "*and only one functional form (one that assumes that the risk of institutionalisation is constant over time)*".

Response

On p17, Lundbeck cite other studies that found increasing hazard of institutionalisation over time. However, all our statistical analysis is performed on the Oxfordshire data, therefore we restrict our focus to this study. We accept the criticism that the fits to time to institutionalisation (Fig 64) and overall survival (Fig 65) could have been improved slightly if we had used more complex 2-parameter models, such as Weibull distributions. However, as stated in our report, we chose the single parameter exponential distribution for simplicity. We believe that the cost-effectiveness results would change only very slightly if we use 2-parameter distributions.

4. The use of the Oxfordshire (Wolstenholme et al. 2002) dataset

Several of the consultees commented that the Wolstenholme study's dataset, of 92 AD patients from Oxfordshire from 1998 to 1999, is not generalisable to the care pathways and resource use of AD patients across England and Wales in 2010.

We are aware of this data source's main limitations for our economic modelling and acknowledge it clearly in our report (Discussion, Section 9.3).

However, all cost-effectiveness modelling exercises have to make use of the best available data for the decision problem at hand, and the need to use individual patient data for modelling Alzheimer's disease is paramount because of the widely appreciated complex and multi-dimensional nature of disease progression. The appraisal committee and consultees should be reminded of the following points which mainly determined our decision to use this Oxfordshire dataset, despite its limitations:

- The reliance on a US patient data rather than UK data to inform the previous economic modelling (the SHTAC-AHEAD model, which informed the previous NICE technology appraisal of these drugs) was regarded as a key limitation of the previous SHTAC-AHEAD model.
- The main outcome of interest – time-to-institutionalisation – for which obtaining and analysing IPD was required could be expected to be quite variable between countries with different health and social care systems.
- At the time that we had to start finalising the structure and input parameters of our 'best supportive care' disease progression model, we had to choose between the two UK datasets available to us: the **Oxfordshire dataset**, and the **LASER-AD** dataset. Use of the Oxfordshire dataset was chosen in preference to the LASER-AD study data mainly because (a) a substantial proportion of patients in the LASER-AD dataset were already receiving treatment with either AChEIs or memantine, and (b) the longer follow-up period of the Oxfordshire dataset (11 years vs. 4.5 years for LASER-AD).

5. The omission of behavioural outcomes from our economic model

A number of the consultees commented that our model does not capture possible behavioural outcomes of the AD drugs.

In Section 9.3 of our report we clearly acknowledge this omission and weakness of our economic modelling.

The main reason that our model does not, as we had originally hoped, include behavioural outcomes is that the IPD study (Wolstenholme et al. 2002) from which we obtained our time-to-institutionalisation regression equations did not contain behavioural outcomes in a form consistent with how the relevant RCTs have assessed this outcome (for example, as NPI scores). (NB. Behavioural outcomes were recorded in the Oxfordshire dataset but only as separate items for mild or severe 'physical aggression', 'aggressive resistance' and 'verbal aggression' – from the Present Behavioural Examination.)

6. Costs by time to institutionalisation in PenTAG model

Criticism

Eisai/Pfizer state on p14: "*The function developed to relate costs to time to institutionalisation does not control for any other individual characteristics, and as evidenced by Figure 70, does not fit the data very well (See Appendix C). There is no indication that PenTAG made any attempt to deal with outliers, deal with non-normally distributed costs or tested alternative functional forms beyond adding polynomial terms to the predictive equation.*

We also note that the cost equations used for mild to moderate disease will predict negative cost for some patients, and as indicated by Figure 70 (note: in the model PenTAG corrects this by artificially assigning a maximum time to institutionalization of 75.5 months), for mild to moderate patients, predict costs lower than those observed in the dataset for individuals whose time to institutionalisation exceeds roughly four years."

Eisai/Pfizer state on p2 "*Equations used to predict costs do not take account of several outlier patients, nor control for patient characteristics, and the functional forms selected do not seem to fit with the observed data. As a result the predictions of the equations do not fit expectations or the data. For example, costs in the immediate time before institutionalisation are greater than those associated with institutionalisation, which is inconsistent with studies such as the Dementia UK report and previous modelling in this area*.".

Response

Figure 70 in our report shows that there is a great deal of variation in the profile of costs as a function of time to institutionalisation. However, in cost-effectiveness analysis, we are concerned with the mean relationship. We have carefully specified an appropriate statistical model, a cubic equation, and uncertainty in this relationship is incorporated in the PSA. We state in our report that terms higher than order 3 were non-significant, therefore omitted from our model.

The criticism that the cubic equations predict negative costs is trivial, because this is extrapolation outside the range of the data, and we constrain costs to be positive.

7. Incorrect Data are extracted from the source paper on patient health utilities

As Eisai/Pfizer point out, the carer-proxy utilities used from the Jonsson et al paper are in fact the average from all three measurement scales: EQ-5D, EQ-5D VAS and QoL-AD. The values used in the PenTAG model and the EQ-5D carer-proxy ratings also presented in the Jonsson et al paper are shown in Table 1. As can be seen when compared with the EQ-5D ratings when both carer-proxy and patient self-ratings are available (the second column of utility values), the utilities are very similar for all MMSE scales except for severe AD (0.33 from the average of scales and 0.4 from where both carer-proxy and patient self-ratings were available). When patient-self ratings are not available (the third column of utility values), the carer-proxy utilities are inconsistent across MMSE. Sensitivity analyses using the EQ-5D values where both carer-proxy and patient self-rating are available and where just carer-proxy are available will be undertaken in the new analyses.

Table 1 Carer-proxy HRQoL from Jonsson et al

MMSE	Values used in PenTAG model	EQ-5D where both carer and patient ratings were available	EQ-5D where just carer-ratings were available
26-30	0.69	0.7	0.5
21-25	0.64	0.65	0.19
15-20	0.5	0.52	0.21
10-14	0.49	0.51	0.39
0-9	0.33	0.4	0.22

8. Modelling 'errors' or anomalies

We thank the reviewers from Eisai/Pfizer and Lundbeck for identifying two modelling errors (a and b below) and anomalies with the secondary structure for calculating utilities and costs in the pre-institutionalisation state (see c below).

a. Referencing to blank cells

There was reference to empty cells in some of the calculations for the secondary structure used to calculate costs and utilities by time to institutionalisation. However, this omission affected, at most, just 1% of the total cohort and as Eisai/Pfizer state it had minimal impact on the ICERs, e.g. the ICER for rivastigmine patches reduced to £60,500 from £61,100 when compared to best supportive care. Note that these ICERs are no longer relevant due to the re-analysis of the modelled treatment effect (see comment 1).

b. Possibility of negative numbers of patients in the Institutionalisation state

This error was an oversight with the PSA which led to negative numbers of individuals in health states. This has been addressed by allowing perfect correlation between coefficients for the prediction equations for time to institutionalisation and overall survival (previously, no correlation was assumed), and by constraining the number of individuals in states to be ≥ 0. A new PSA has been undertaken for the new base case analyses for mild to moderate AD and moderate to severe AD, and will be presented with the full re-analysis of the decision model in line with response 1 above.

c. Anomalies with calculation of costs and QALYs in the pre-institutionalisation state

Eisai/Pfizer and Lundbeck pointed out that average costs and utilities per patient are constant across the model time horizon. This is correct. In the PenTAG model, pre-institutionalisation utilities and costs are modelled based on time to institutionalisation. To calculate the number of individuals at specific time periods prior to being institutionalised, the exponential model predicting pre-institution state occupancy is used. This allows utilities based on MMSE, and costs based on time to institutionalisation to be distributed across the population within the pre-institutional state. Since an exponential model is used to predict time to institutionalization, the proportions of individuals at specific time periods prior to being institutionalised are constant for the model time horizon. Therefore, the monthly utilities and costs per patient in the pre-institutionalised state are constant for the model time horizon. Calculation of the pre-institutionalised costs and utilities in this way, allows for a treatment effect to impact on both the costs and utilities. The PenTAG model estimates lower pre-institutionalised costs and higher pre-institutionalised utilities per patient for the treated cohorts compared to the best supportive care cohorts.

Related to this point are further comments from Lundbeck*:*
"PenTAG implemented a modeling structure that purports to allow for diminishing utility and rising costs over time thereby addressing limitations of the previous model. However, an inspection of the model reveals that the calculations do not implement these changes in cost and utility and that the calculations used in fact generate values that vary considerably from the source data. It should also be noted that in the model utilities vary with age although no such relationship was postulated. The secondary structure also introduces an inconsistency in the key input – the starting MMSE."

We acknowledge that since age is a statistically significant predictor of time to institutionalisation and that to model decreasing utility and increasing cost in the pre-institutionalisation state, the equation for time to inst has been used, therefore inducing an unintentional relationship between utility/costs and age.

As Lundbeck point out there is an inconsistency between the starting MMSE for a cohort and the average MMSE calculated from the PenTAG model, which impacts upon the average utility calculated for the cohort. This relates to the inconsistency between the equations raised by Eisai/Pfizer and addressed in point 1 of our response to consultee's comments. The survival equation from the Oxfordshire dataset to predict time until the end of pre-institutionalisation is based on MMSE, Barthel and age. We then take time prior to the end of pre-

institutionalisation and input this into the repeated measures equation from the Oxfordshire dataset to calculate a corresponding MMSE (MMSE = 8.34 + 4.17t). It is therefore not surprising that this inconsistency exists as two different equations are applied to the data and especially as for the survival analysis data from all individuals are included, while for the repeated measures analysis only data for those individuals becoming institutionalised are included. Thus, the average utility for the pre-institutionalisation state in the best supportive care cohort is estimated to be 0.521 from the PenTAG model for mild to moderate AD, and 0.512 for moderate to severe AD. The data source indicates that for a MMSE of 17 (mild to moderate) the utility is 0.5 and for a MMSE of 11.7 (moderate to severe) the utility is 0.49. It has not been possible to fully address this issue for these re-analyses, however as it is the incremental QALYS and costs that are of interest, we do not believe this would have much of an impact on the new, updated ICERs. Given these overestimates of utility compared to the data source, we believe that the impact of this is to give a larger benefit to the active treatments (since the delay to institutionalisation is estimated to be at a slightly greater utility than the evidence indicates).

9. Mapping from ACDS-ADL to Barthel indices performed by PenTAG

Criticism # 1

Lundbeck state on p23;

"*In the mapping approach undertaken by PenTAG, when no single question from the ADCS-ADL19 matched a given Barthel scale item, the rescaled total score on the ADCS-ADL19 (i.e. the sum of the basic and instrumental activities) was used as a proxy for the Barthel scale item. This implicitly assumes that the missing Barthel scale item ("missing" in the sense "no equivalent in ADCS-ADL19") would have been an "average" item of the ADCS-ADL19, if it had been measured in this scale. The term "average" refers to the ranking of ADLs, from the most simple (i.e. least impaired, with highest score) to the most difficult (i.e. most impaired, with lowest score). As the ADC-ADL19 is made of 6 basic activities (the first six items) and 13 instrumental activities, the above-mentioned missing Barthel scale item is then expected to be among the simplest ADLs and have a higher score compared to an "average" item of ADCS-ADL19*.

This obviously results in an underestimation of this missing Barthel scale item, which then cascades into an underestimation of the total Barthel index score. This ultimately leads to an underestimation of the treatment benefit estimated from this mapping approach and therefore an underestimation of the time to institutionalisation."

Response # 1

We thank Lundbeck for taking the time to investigate the mapping in such detail, however, it is simply incorrect that "*the above-mentioned missing Barthel scale item is then expected to be among the simplest ADLs*". Therefore this accusation is dismissed. Consider the Barthel question concerning "Transfer". We found no questions on the ADCS-ADL19 (severe scale) which directly deals with this. The Barthel score for the "Transfer" question was therefore estimated as the maximum possible score for this question on the Barthel scale multiplied by the ratio of (the total score on the ADCS-ADL19 questionnaire / the maximum possible score on the ADCS-ADL19 questionnaire overall).

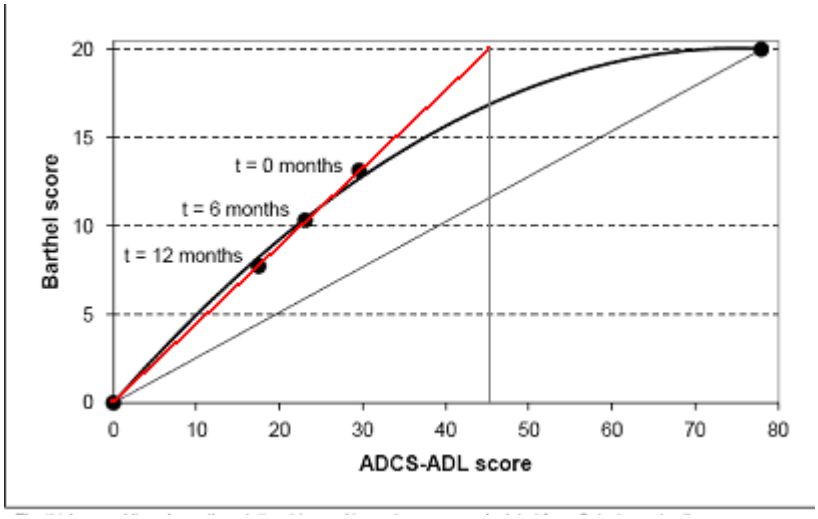
We also note that of the 10 questions on the Barthel scale, the score on only 2 ("Transfer" and "Stairs") were estimated in this manner. Therefore, if their criticism were valid (which it is not), it would only have a very trivial impact on the mapping algorithm.

Criticism # 2

Lundbeck state on p23;

"*Also, as shown in figure 61, the relationship between the Barthel index and the ADCS-ADL score seems very close to linear if the constraint on maximum scores is omitted. Following a linear assumption, the maximum Barthel index score would then translate into a score below the maximum of the ADCS-ADL (between 45 and 50 on a scale of 78). This is consistent with the different contents of the scales considering that the Barthel index does not include instrumental activities of daily living. As instrumental ADLs are more complicated and deteriorate faster than the basic ADLs, a maximum score on the basic activities (Barthel index) does not guarantee a maximum score on all activities (ADCS-ADL).*"

Figure 6



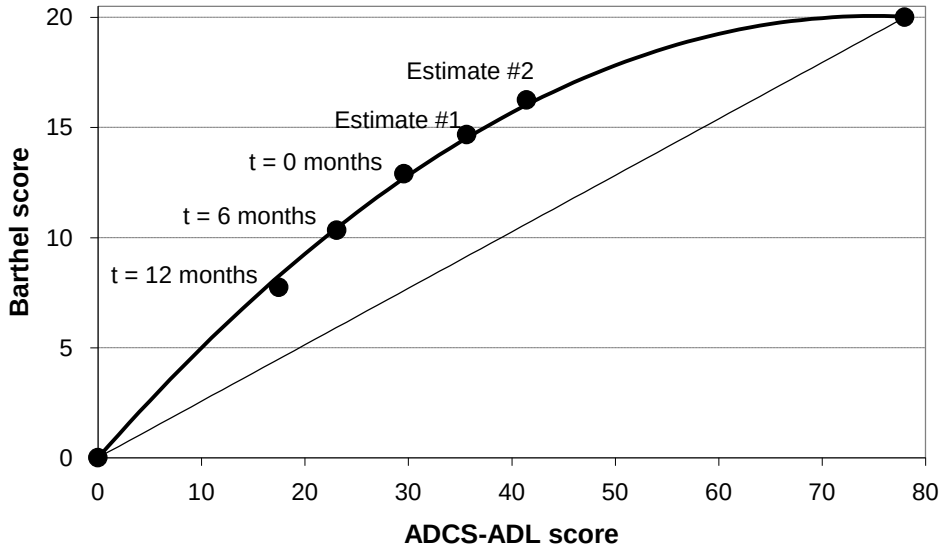
Response # 2

This is an interesting suggestion. The mapping must pass through (77,20), the maximum scores in the ADCS-ADL and Barthel scales respectively, because if a patient scores the maximum of 77 on the ADCS-ADL scale, given that the Barthel questions are effectively a subset of the ADCS-ADL questions, then the patient also scores the maximum of 20 on the Barthel scale. The important question is then what is the shape of the mapping between the t = 0 months point and (77,20) ? This portion of the mapping is particularly important because it encompasses the initial ADCS-ADL scores for the drugs in the PenTAG model. Indeed, the portion of the mapping from (0,0) to the t=12 months points is not used in our model. From our calculations, we estimate that if a patient scores the maximum of 20 on the Barthel scale, the minimum possible scores on the ADCS-ADL scale is approximately 34. This does not, of course, imply that the

mapping passes through the point (34, 20). Instead, the mapping must lie beneath this point.

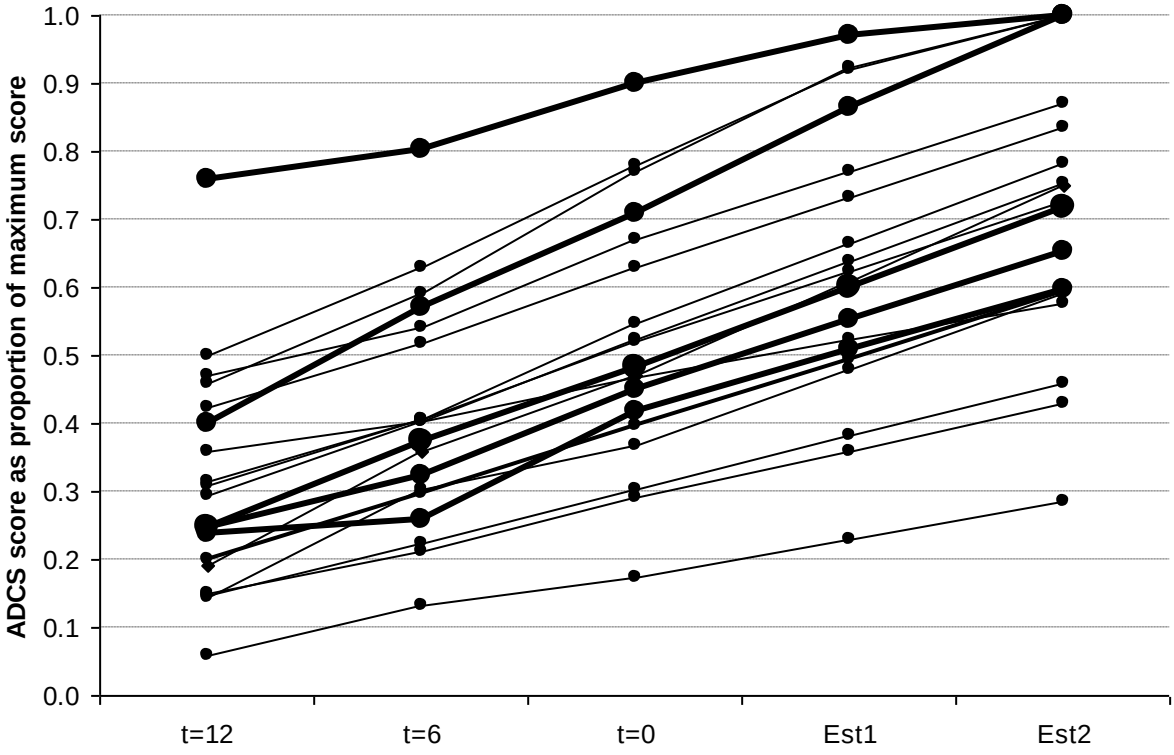
We estimate the positions of two further points on this graph as shown in the graph below, Figure 7. These points clearly lie very close to our original mapping. Therefore, we maintain that our original mapping is appropriate, and that the mapping as suggested by Lundbeck, shown in their graph above, is inappropriate.

Figure 7



We derived these two points as follows. The graph below, Figure 8, shows the ADCS-ADL-severe scores for each of the 19 questions as a proportion of the total question score for each of the time points t=0, 6 months and 12 months from Galasko et al (2005). Next, we extrapolated each of these lines independently for each of the questions, ensuring that the proportions are capped at 1. These are shown by Est 1 and Est 2 in the graph below. The first six questions of the ADCS-ADL-severe questionnaire are shown as thick lines, and the remaining 13 questions as thin lines.

Figure 9



The estimated proportions were then simply converted to absolute scores for each of the 19 questions. Next, the estimated scores on the Barthel index and the ADCS-ADL-23 index were independently calculated, in exactly the same way as for the actual 3 data points, and as described in our original report.

Criticism # 3

Eisai/Pfizer state on p2; "*However, the mapping was based on 3 data points only, which do not contain over 60% of the range in possible ADCS-ADL scores. Reasonable alternative mapping equations produce dramatically different results from the model. For example, assuming a constant relationship between the two scales reduces the cost per QALY for donepezil by approximately £60,000 to £23,240*"

Eisai/Pfizer state on p15; "*The equations used to transform ADCS-ADL scores to Barthel scores are almost completely arbitrary*….."

Eisai/Pfizer state on p15; "*The function used by PenTAG also results in some implausible predictions. For example, the function predicts that improvements in ADCS-ADL are associated with deterioration in Barthel scores for ADCS-ADL scores above 75.*

An improvement in ADCS-ADL from 60 to 70 (10 points) is equivalent to a +0.66 change in Barthel score. From 70 to 80, it is equivalent to a change of -0.66."

Response #3

Each of these 3 data points are the average scores from 145 patients. Therefore we can attach a large degree of certainty to the 3 data points.

The quadratic mapping equation is clearly based on much careful analysis (see p276-8 our report). Therefore, we totally reject the allegation that it is "*completely arbitrary*".

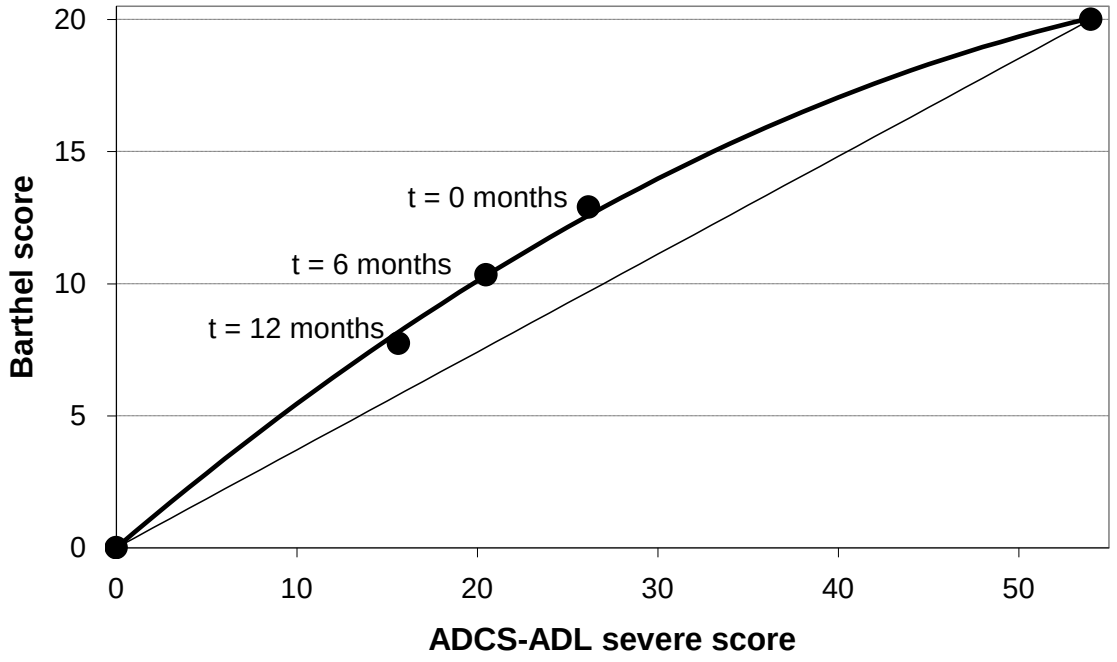
Eisai/Pfizer suggest a simple linear mapping. This is clearly inferior to our method, where we have carefully derived a quadratic mapping based on much data.

The allegation of implausible predictions from the quadratic mapping is extremely trivial. First, we correct the statement that an increase in ADL from 70 to 80 equates to a decrease in Barthel of 0.66. In fact, the change in Barthel is virtually zero. Second, it is true that our mapping predicts a decrease in Barthel for ADL above 75. However, the maximum possible ADL is 77, and for ADL scores above 70, the Barthel is virtually unchanged. This is perfectly plausible.

10. Memantine data was inappropriately handled when assessing the effect on the Barthel index.

In our report, we assumed that the treatment effect for memantine of 1.408 was measured on the ADCS-ADL-23 scale, whereas it is, as Lundbeck say, actually measured on the ADCS-ADL-severe scale. We have corrected our error as follows. In our original submission, we derived a mapping from the ADCS-ADL-23 scale to the Barthel scale (Figure 61 in our report). Using the same data from Galasko et al (2005), we repeated this procedure to obtain a quadratic mapping from the ADCS-ADL-severe scale to the Barthel scale, see the Figure 10 below.

Figure 10



The mapping is: Barthel = 0.5835 (ADCS-ADL-severe) - 0.0039 (ADCS-ADL-severe)^2. As for the previous mapping, uncertainty in this relationship was incorporated in the PSA using the Cholesky matrix from the regression analysis. Previously, we estimated the treatment effect of memantine to be 0.32 on the Barthel scale, whereas when corrected, this increases to 0.40.

13. QALYs reducing when cognitive or functional status improve

Eisai/Pfizer noted that in the PenTAG model an improvement in MMSE or Barthel leads to negative incremental QALYs for donepezil compared with best supportive care. As they state, this is related to the prediction equations for time to institutionalisation on which the treatment effect is assumed, in particular, the uncertainty associated with the estimated coefficients. This issue is no longer relevant in the new analyses for incorporating the treatment effect (see comment 1 above).

14. Cost of AD drugs in PenTAG model

Rivastigmine cost: The Novartis Consultee comments question the daily drug cost for Rivastigmine capsules which we have used in our model. Our weighted average daily cost of taking Rivastigmine capsules assumes that 70% of patients are on 9mg per day and 30% are on the maximum daily dose of 12mg per day (based on the approximate estimated dose levels of the included RCTs).

Novartis are correct that the PenTAG calculations of the cost of 9mg daily dose (£3.56) are calculated as taking the 3mg capsule 3 times per day. If it was taken as 2 × 4.5mg capsules then the cost of taking 9mg daily would be £2.38, as Novartis state. In one of the two RCTs from which we obtained our effectiveness estimates for Rivastigmine

(Feldman & Lane, 2007; ref. no. 132 in the report) the 9mg dose was achieved as three 3mg capsules per day.

Therefore, achieving the 9mg dose with 2 rather than 3 capsules per day would reduce the estimated daily cost of Rivastigmine capsules in our model from £3.21 per day to £2.38 per day (a 26% reduction in drug costs).

Updating of monthly drug costs for rivastigmine capsules from £98 to £72 reduces the total costs from £70,892 to £70,686, resulting in an ICER of £69,700 compared to best supportive care. However rivastigmine capsules remain dominated for the incremental ICERs. Note that these ICERs are no longer relevant due to the re-analysis of the modelled treatment effect (see comment 1).

Donepezil cost: The Eisai/Pfizer Consultee comments indicate that the daily cost of donepezil was reduced in January 2010 by 5.8%, to £3.00 per day for 10mg. The daily drug cost in the PenTAG model is £3.18, which is directly calculated from the most current edition of the British National Formulary (BNF) which was available when accessed on the 2nd of March 2010. For pharmaceutical products the NICE *Guide to the methods of technology appraisal* states that the "public list price" should be used, and the most current edition of the BNF is the standard source in most if not all technology assessments for such list prices.

The daily price difference of £0.18 would make a very small difference to the estimated ICERs.

Clinical effectiveness systematic review

Lundbeck

1. P 35 PenTAG limited their assessment of memantine to moderately-severe to severe AD patients only. Furthermore, their review excluded patients who were on a stable dose of AChEIs.

We did not limit our assessment to moderately-severe to severe AD people. We included Porsteinsson et al. 2008 whose population MMSE range was up to 20.37. Section 4.8. This study also included patients on stable doses of AChEIs.

2. P35 The fact that PenTAG undertook a restricted review of the evidence was not clearly stated.

The limitations of the inclusion criteria in the PenTAG systematic review of clinical effectiveness are clearly stated in Sections 1.1, 3, and 4.1.2.

3. **P35 and p36.** Lundbeck criticise us here for failure to include their meta-analysis of six studies. They go on to say that we only included two of the six studies.

Lundbeck's meta-analysis includes the following studies: Peskind et al 2006, Bakchine and Loft, 2007, Porsteinsson et al. 2008, Reisburg et al 2003, van Dyck et al. 2007 and Tariot et al. 2004.

We screened this meta-analysis and found that two of the studies did not meet the inclusion criteria for the scope of this report. These were were Peskind et al. 2006 and Bakchine and Loft, 2007 which were both excluded because they included people with mild AD. Therefore, it was not possible to include the meta-analysis per say because the results would have been skewed by the participants with mild AD. The other **four** studies were included in our systematic review.

4. **Study MEM-MD-12: Porsteinsson et al, 2008 includes mild AD patients who fall outside the memantine indication and should be excluded.**

This is the same study that Lundbeck criticised us for excluding above (although we included it). The upper range of the MMSE scores for the participants of this study was 20.37. We took the view that as this was only minimally over the threshold of 20 we would include this study; particularly as it was the only new evidence we found for combination therapy.

Eisai/Pfizer

Eisai/Pfizer claim, on p15 of their comments, that Donepezil has significant benefit on NPI. This claim is based on a meta-analysis including participants from outside the licensed MMSE range. It is noteworthy that, although it is an included study in the meta-analysis in question, the outcome pooling does not include the AD2000 study; this is important, as AD2000 has a strong influence on the null result of PenTAG's pooled analysis (see TAR Figures 13 & 14).

Shire Pharmaceuticals

In their second section, on p2 of their comments, Shire asserts that we have 'concentrated on the small amount of post 2004 data'. This is not the case, in fact we have included the trials that were included in the 2004 TAR and, where the data permit, pooled the results of these trials with the new trials to provide revised pooled estimates of the effectiveness of all the drugs.