HEALTH TECHNOLOGY APPRAISAL: NICE Health Technology Appraisal - Assessment Report Donepezil, galantamine, rivastigmine and memantine for Alzheimer's disease (Review)	
TO: NICE	**FROM: NHS Quality Improvement Scotland 10 August 2010**

Provided by:
[redacted]
Consultant Psychiatrist
3 August 2010

Although the report is lengthy and at times overly repetitive the conclusions of the report appear to be broadly in line with clinical experience. In addition, the relatively low number of randomised controlled trials, published since the last review, are not surprising in a market where the companies concerned have no real competitors in terms of alternative treatment strategies. The lack of focus on quality of life for people with dementia and their carers and the absence of long term outcome measures is particularly disappointing.

The report refers to ongoing studies including DOMINO-AD. The authors of the report continually refer to this study as being applicable to those in whom cholinesterase inhibitor treatment has failed but this is incorrect. The study population is people who are at the lower end of the NICE guidance recommendations for treatment with cholinesterase inhibitors but such patients may well have responded to cholinesterase inhibitors at an earlier point in their treatment. The study, though likely to be very underpowered, will provide some clues as to what should be done with patients who are deteriorating as their illness progresses although a definitive answer is unlikely to be found.

The report highlights more recent high quality studies for Rivastigmine and Galantamine. I think this is not surprising given that these compounds post dated the launch of Donepezil. New studies continue to be concentrated on patients with a mean MMSE in the "moderate" range and definitive studies in mild dementia are still lacking. Extrapolation from moderate dementia or from treatment of MCI is not appropriate.

The report refers to the alternative to cholinesterase inhibitors as being "best supportive care". However this is not defined. It is known that the quality and quantity of care provided to people with dementia and their carers varies widely across the country and may bear little relationship to the care provided

to patients who are enrolled in drug studies. In fact evidence for good quality community care being effective is very limited. An American study (The Medicare Alzheimer's Disease Demonstration Evaluation) would suggest that community care purchased at a level similar to the cost of cholinesterase inhibitor was ineffective in reducing institutional care or indeed most other outcome measures.

I agree that many of the RCTs to which the reports authors refer add little to the overall literature. The better quality studies not surprisingly adjust pre-existing results across the domains of cognition, function, behaviour and global outcome to demonstrate that the efficacy of Donepezil, Rivastigmine and Galantamine are broadly similar with substantial overlaps in efficacy being plotted in each case. Indeed on head-to-head comparisons no consistent results emerge in keeping with the clinical impression that the efficacy of the three compounds is similar.

The same pattern emerges when considering cost effectiveness data. Again this is presented as at times showing one compound to be better than another though perhaps surprisingly Donepezil fairs least well.

The lack of good quality data supporting the use of Memantine could be seen as surprising in view of previous NICE guidance but this drug is widely available in Europe and the US and I suspect that the company have little motivation to address a recommendation applicable to a relatively small market. The data from the new study has a negative effect on the previous data with the result that Memantine is shown to have very little effect by comparison to placebo in any of the domains. There is some suggestion of improved effect where patients have agitation, aggression and psychosis but even here cost effectiveness data is not supportive.

Data on combinations of cholinesterase inhibitors and Memantine are also disappointing. However more information may be forthcoming when DOMINO-AD results are published.

The cost effectiveness section of the report is largely outwith my expertise but in view of the history of disagreement between Eisai/Pfizer and assessors engaged by NICE a great deal of the cost/ effectiveness section is spent rebutting arguments put forward by industry and laboriously explaining the rationale for their own model is far from unexpected. The reports authors are to be commended for trying to find UK data on which to base their economic model though I cannot see that their decision to base their modelling on a study of 92 patients in Oxfordshire and another relatively small longitudinal study (Laser-AD) is going to go unchallenged by industry.

Perhaps because of the different models and different underlying assumptions used there is a huge gulf between the ICER values generated by Eisai/Pfizer and those generated by the reports authors. I cannot see how these values can be reconciled.

In conclusion I think that the report is unlikely to change the conclusions of TA111. Current NICE guidance is not in keeping with clinical experience where one is concerned with individuals from the general population rather than a selected population who participate in drug trials particularly since it is not clear that recruits from the UK match the demography of the general population.

The patent on cholinesterase inhibitors will lapse before the next NICE Review is due and it is almost certain that prescribing patterns will change once this happens. It is important that this is taken into account during the current consultation. It is disappointing that after 12 years of their use much of the debate on efficacy rests on the interpretation of drug trials which were recruiting in the mid to late 1990s. Key unanswered questions about long term effectiveness, quality of life and maintaining people in the community remain answered. The contribution of cholinesterase inhibitor treatment to reduce the level of antipsychotic usage also needs to be established given concerns about iatrogenic mortality associated with antipsychotic use in people with dementia. It is unlikely that industry led trials will now take place.

In Scotland NICE guidance is considered to be too narrow and is seldom strictly applied. The advent of Free Personal Care has also changed the profile of community services and this may alter the question on what interventions are most appropriate for people with dementia and their carers. The clinical benefits of cholinesterase inhibitors are known to be modest but the treatment of dementia is characterised by interventions with small effect sizes and part of the skill in treating people is to try to have additive effects. It is rare for cholinesterase inhibitors to be the only intervention a person receives.

In Scotland there is limited experience with the use of Memantine. Postcode prescribing is more marked than with cholinesterase inhibitors. As a consequence very mixed results are reported from clinical practice. There is no doubt that some people respond well particularly those in whom agitation and aggression are markedly reduced. In many other cases, however, response is minimal or increased agitation leads to drug withdrawal. Although some Health Boards have a system for authorising its prescription on an “exceptional” basis there is no doubt that some clinicians feel it is impossible to make a successful argument to bodies who are determined that Memantine should not be used. The current limitations on its use are too strict but efficacy data would not support its routine use. Some middle ground needs to be found for the wording of the forthcoming NICE recommendation.

- 0 -

Provided by:
[redacted]

Background of appraiser

I have been invited to serve as a clinical expert for Quality Improvement Scotland in the appraisal of the above report prepared by the Peninsula Technology Assessment Group (penTAG) as the University of Exeter under the leadership of Professor Chris Hyde, Professor of Public Health at the University of Exeter. I am qualified as a specialist in Old Age Psychiatry and General Adult Psychiatry. I have focussed on Alzheimer's disease throughout my research career, initially on the discovery that the neurofibrillary tangles of Alzheimer's disease are composed of the protein Tau, and from there to the development of a treatment aimed at arresting the progression of AD as a potential disease-modifying treatment based on stopping the underlying aggregation of Tau protein. This research has now reached the stage of having completed a large UK-based Phase II clinical trial in 321 subjects providing initial evidence of arrest of rate of progression of AD by approximately 80% over 12 months as measured by rate of change in ADAS-cog and supported by neuroimaging. We expect to be able to initiate an international confirmatory Phase III trial in January 2011. The vehicle for this research and development since 2002 is a spin-out company of the University of Aberdeen, TauRx Pharmaceuticals, of which I am Chairman, in which I have a financial interest.

Further assistance in preparing this review was obtained from my son, Dr. Damon Wischik, who is a Royal Society Research Fellow who lectures at UCL in mathematical modelling of networks, and whose general research field is applied probability.

PenTAG review of efficacy

The stated remit for PenTAG report was "up-date the evidence used to inform the last NICE guidance on donepezil, galantamine, rivastigmine and memantine for the treatment of Alzheimer's disease, particularly as laid out in the report by the Southampton Health Technology Assessments Centre (SHTAC). In general they considered evidence up to 2004, and this is the start date we have used for this report (page 365)."

The PenTAG report did not undertake a re-examination of all the available evidence from scratch, but did undertake a review of the efficacy of the AD treatments in light of 17 new randomised control clinical trials (RCTs) of varying quality which have become available since the SHTAC review in 2004. Although they undertook a detailed review of only the newly emergent studies, their final estimates of efficacy for each of the treatments under review incorporated RCT data from the older and as well as more recent reports. The net result is a synthesis of data from the new trials and from the earlier SHTAC review using the statistical technique of random-effects meta-analysis to derive new estimates of treatment efficacy with

respect to placebo at approximately 12 and 24 weeks for the key outcome measures in the domains of cognition (ADAS-cog and MMSE), function (ADCS-ADL), behaviour and mood (NPI), global function (CIBIC-plus and CDR). From this analysis they derived weighted mean differences (WMD's) as measures of effect size which were used in the ecometric modelling and standardised mean differences (SMDs) to determine the overall statistical significance of pooled efficacy estimates from multiple sources.
Broadly this part of the review was extremely thorough, capturing data in a highly systematic manner, and using sophisticated and appropriate analysis techniques to derive pooled overall estimates of efficacy.
It is interesting to reflect, in light of the general complaints made in the PenTAG report regarding the inadequacy of the available data base, that the efficacy data base which informs this review is considerable. Estimating at today's clinical trial costs, the expenditure required to achieve the information which is now available for review is approximately £600 million.
Subjects 2004 2010
Donepezil 4,465 4,256 209
Galantamine 4,653 3,324 1,329
Rivastigmine 3,894 1,940 1,954
Memantine 602 252 350
Comparative 1,421 259 1,620
Total 15,035 10,031 5,004
As can be seen, the new data is weighted to rivastigmine and galantamine. This somewhat shifts the report's view of the clinical merits of rivistigmine and galantamine relative to donepezil, and the new evidence supporting the efficacy of memantine remains equivocal. There is substantially more head-to-head comparison data, which broadly confirms that these treatments should all be considered as a class, as there is little to distinguish them.
Broadly speaking, the PenTAG review confirms the generally prevalent impression in the field that the treatments under review are effective in the domains of improving cognition, function, behaviour and global impact. Nevertheless, the report highlights the important gaps in knowledge which remain, particularly as this effects long-term outcomes, impact on quality of life, impact on carers and time to institution. The PenTAG review correctly highlights that the available efficacy data is generally restricted to the time horizon of 6 months, and the report also correctly highlights the difficulty in extrapolating from these 6-month effects to an economic analysis which fundamentally requires a minimum of a 5-year horizon to be meaningful (let alone the 20-year horizon implicit in the PenTAG model).

PenTAG economic modelling

The economic part of the report consists of two parts, a critique of cost-utility assessments submitted to NICE by Lundbeck (memantine) and Pfizer/Esai (donepezil) and the creation of a new PenTAG cost-utility model.
The PenTAG group criticises the methodology adopted by Pfizer/Esai as well as the underlying assumptions. In respect of the Lundbeck model, the critique is limited to the underlying assumptions. There is little in principle to

distinguish an approach based on a Discrete Event Simulator (DES, model provided by Pfizer/Esai) and the Markov-based approach adopted by PenTAG. The PenTAG review complains that the Pfizer DES model is not pure, but it can equally be said that the PenTAG Markov model is likewise not pure, in that transition rates between states depend on time since the start of treatment as well as on the present state. Therefore, the PenTAG model is better characterised as a “Markov-inspired DES”. It is in any case possible to construct a one-to-one mapping of DES and Markov models, and the difference resides purely in precise specification of assumptions. In the present circumstances, a Markov-type model is probably more suited to the purpose in that the operation of the underlying parametrisation and assumptions can be defined in a more transparent manner. The principal difficulty is that a Markov-based model needs to be somewhat more sophisticated than that provided by the PenTAG group, as discussed further below. Clearly, the PenTAG group had difficulty in dealing with the generality of the Pfizer/Esai DES implementation, and so its potential benefits appear to have been lost for the purpose to hand. As it stands, the Pfizer/Esai model is largely dismissed because of the property identified in Table 100 whereby there is an inexplicable relationship between between cycle time in the model and systematic change in QALY benefits and cost savings.

The real significance of this, beyond dark hints that “it may be due to an error in the model logic”, is unclear, and could surely have been resolved by a simple exchange between competent probabilists on both sides.

The fundamental weakness of the PenTAG model stems from strategic decisions outlined on pages 259 – 261. The most important weakness was the decision to model transition rates derived from 6m RCT data to the whole subsequent period, in particular the period after initiation of treatment and prior to institutionalisation, without taking proper account of the underlying nature of the treatments in question or the impact on efficacy of discontinuation of treatment.

This review will therefore concentrate on a critique of certain fundamental aspects of the PenTAG model in the hope that with some further input, the model can be made better fit for purpose. This is done in the absence of access to the model itself, which would permit gaining a better sense of its sensitivity to various changes. Although a sensitivity analysis is presented by the PenTAG group, it is hard to get a sense of the actual dynamic range of the model without having it to experiment with.

The essence of the model is presented on pages 283 and 285, whereby a value λPre-inst (the rate of state transition to institutionalisation) to be used in the Markov model is calculated as an exponential function of MMSE, ADL and age at study entry. A corresponding rate term λOS (the rate of state transition to death) is calculated as an exponential function of MMSE, Barthel and age at study entry. These are the two rate terms in the model expressing rates of transition between start of treatment (or non-treatment) to the final state of death, via institutionalisation if this precedes death. The constant terms in the λ equations were estimated by fitting the functions to a data set derived from a prevalent cohort of 92 patients followed for 11 years between 1988/9 to 1997/8 in a UK study by Wolstenholme et al.

(*Estimating the relationship between disease progression and cost of care in dementia*. Brit J

Psych 2002; 181:36-42) who were on average 4.9 years post diagnosis at point of entry into the study. The rate terms are applied at monthly iterations of the model to calculate time to either institutionalisation or death for any given individual, where mean time is given straightforwardly by 1/λ. The model is applied at 3 levels of severity and 3 age groups with corresponding distributions making up the theoretical starting cohort. Time in a given state is then transformed into a utility function. Cost likewise is calculated as a function of time according to the two equations (12) and (13) given on page 302. The rest of the report consists simply in a discussion of the derivation of the terms which enter these equations, and the sensitivity analysis consists in making small plausible tweaks to the assumed values.

The fundamental structural difficulty of the PenTAG model as it stands is that there are inbuilt into it two different rate terms for treated and untreated subjects which survive throughout the time horizon of the study. This is in contrast to the approach discussed in Appendix 16, and Figure 67 in particular, which was apparently not followed during model development. Figure 67 expresses well the generally accepted understanding in the field that symptomatic treatments of the AChEI class and memantine produce a short initial treatment response followed by longer-term progression at the same rate in both treated and untreated groups.

This fundamentally important feature is not captured in the PenTAG model. The PenTAG model implies a structural difference in the rate of disease progression once a treatment period has occurred and irrespective of discontinuation during the period of treatment
(because of the ITT analysis assumption). It appears to be implicit in the thinking of the authors that symptomatic treatments do indeed have an effect on disease progression, for example in the Conclusions on page 38 (“this update systematic review continues to suggest that there is clinical benefit from AChEI’s in alleviating symptoms and *controlling disease progression* in Alzheimer’s disease”).

In other words, there is an implicit assumption in the underlying structure of PenTAG model as it stands that treatment with a symptomatic drug produces a persisting difference in rate of disease progression both in the discussion and in the mathematics of the model itself. But this certainly does not correspond either to the mechanism of action of these treatments or the general understanding of their clinical impact. It is generally accepted that the period of benefit is approximately 6 months, after which patients continue to decline at the non-treated rate, despite continuing to take treatment. Of many references which could be adduced to this effect, one can take for example the EMA Guidance CPMP/EWP/553/95 Rev. 1 **Guidelines on medicinal products for the treatment of Alzheimer’s disease and other dementias** which states: *Based on efficacy and safety data several drugs have been approved for symptomatic improvement of dementia of the Alzheimer Type and one for the symptomatic improvement of dementia associated with Parkinson’s Disease. However, established treatment effects must be considered as modes. Randomized clinical trials in other subtypes of dementia (e.g. vascular dementia) have not been able to demonstrate clinically relevant symptomatic improvement nor was it yet possible to establish disease modifying effects in any dementia*

syndrome or its subtypes… Up to now no clinical trial has led to a successful claim of disease modification in dementing conditions. For regulatory purposes a disease modifying effect will be considered when the pharmacologic treatment delays the underlying pathological or pathophysiological disease processes and when this is accompanied by an improvement of clinical signs and symptoms of the dementing condition.
It is difficult from the report to calculate the magnitude of this falsely implied disease modifying effect. The 6-month MMSE effect sizes that enter the model for mild/moderate AD are 1.24 (donepezil), 1.13 (galantamine), 1.02 (rivastigmine caps), 1.10 (rivastigmine patches) and 0.7 (memantine). The corresponding placebo decline rates have not been estimated systematically by comparison, but using 12-month data from Gold (*Study design factors and patient demographics and their effect on the decline of placebo-treated subjects in randomized clinical trials in Alzheimer's disease*. J Clin Psychiatry 2007; 68:430-438) it can be calculated that the expected untreated placebo decline rate would be 2.4 MMSE units in 12-months or 1.2 units in 6 months. In other words, the implied treatment effect approximates to a complete arrest of disease progression at least for the first 6 months. It is then difficult to understand how it is that given the 6-month head-start for the treated group and the implied disease modifying effect due to assignment of different rates of progression by using different rates for treated and untreated subjects, the final net calculated benefit is only 10 – 12 days delay in time to institutionalisation for mild/moderate subjects in the base case relative to nontreatment. It is not surprising that the calculation of the final incremental cost effectiveness ratio (ICER) should produce a cost in excess of £60,000 per QALY. A further issue is the impact of discontinuation. The PenTag model assumes that treatment benefits persist after treatment has ceased (eg, page 375: “it is assumed in the PenTAG model that treatment benefits remain after treatment has ceased”). This is contrary to clinical experience or general understanding in the field. The discussion of the impact of discontinuation appear on pages 279-280. The approach taken is based on the assertion of the priority of the ITT analysis, while acknowledging the potential for LOCF (last observation carried forward) to inflate effect size due to biased withdrawal of subjects on active treatment avoiding side effects. On page 332, the report acknowledges the important impact of discontinuation, but reaches the paradoxical conclusion that: “higher estimates [of discontinuation rates] lead to fewer costs and greater net benefit associated with AChEIs”. In other words, the smaller the number of patients continuing with treatment the greater the calculated benefit!

Even so, the PenTAG model is favourable to AChEI treatments in this respect, since it assumes proportion discontinuing is linear with respect to time, such that approximately 55% of subjects are still receiving treatment after 12 months in the base case. However, the discontinuation function is non-linear, and there may be as few of 15% of subjects still receiving treatment after 12 months (Figure 1), and overall median time on treatment may be as little as 120 days.
Figure 1.

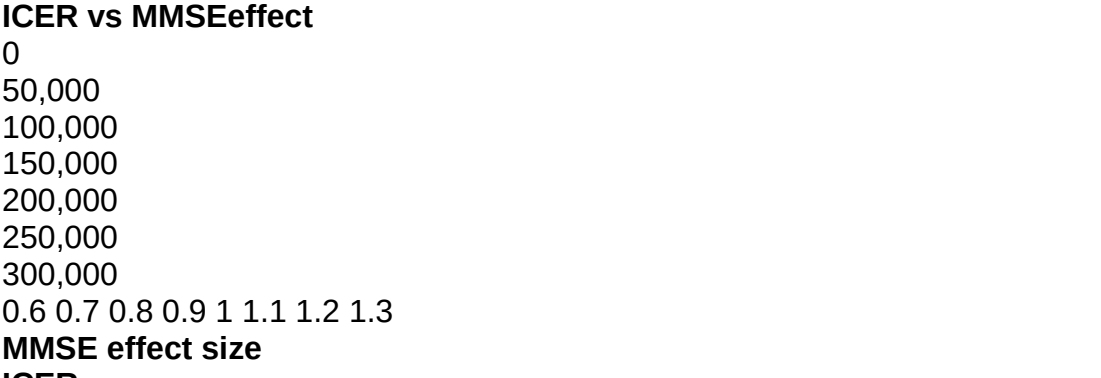
Overall assessment

Overall, the PenTAG report is highly competent and thorough. The team has gone some way to correcting at least some of the defects of the earlier SHTAC model, and the report provides a comprehensive analysis of the field, and is reasonably transparent in its methodology. There are many aspects that one could discuss, some of which the PenTAG report also discusses, such as the limitations of the Wolstenholme data set as a basis, the volatility of the estimates, the uncertainty of the underlying parameter assumptions, etc. A further point which could be added that the conversion between MMSE and ADAS-cog is an important building block in derivation of the fundamental assumptions. However, we found that whereas static comparisons of MMSE and ADAS-cog were correlated (via a quadratic function) there was almost no correlation between *change in* ADAS-cog an *change in* MMSE. This suggests that treatments impact on different aspects of these scales in ways that are not understood, making it very difficult to infer a change in one from a change in the other, despite the fact this type of conversion is often applied in practice. The most important defect in the PenTAG model, in the opinion of this reviewer, is that there was a fundamental strategic error in model development. This could be rectified with perhaps 6 weeks of further input from a competent probabilist working with the team. In essence, the operation of the Markov model that has been developed does not correspond to the generally understood effect of the symptomatic treatments currently available. This mismatch must undermine the credibility of the model which will in any case be severely challenged by a clinical community who have a far poorer grasp of the underlying mathematics. It is not enough to point to the underlying uncertainty and volatility of the model. The fact is that the model will be used as an authoritative basis for further discussion, and needs to be developed further to serve this important purpose adequately.

Because of this strategic modelling defect, the overall conclusions are more favourable to the presently available symptomatic treatments than would be warranted under the terms of reference of the analysis. That is, it would be expected that a more sophisticated Markovinspired approach that did implement a uniform rate of decline after termination of treatment irrespective of treatment (ie the treatments are fundamentally only symptomatic not disease modifying) would establish even higher ICER values than those of the PenTAG report, and confirm further that the present symptomatic treatments are not formally cost-effective within the framework of the terms of reference set by NICE.

Because the model has not been available to the present reviewer to test how it operates, there is a *prima facie* concern that the model could not in principle demonstrate cost effectiveness for any treatment in Alzheimer's disease, even if it were able to arrest disease progression. This concern is illustrated in Figure 2, using values from the PenTAG report. The figure may be misleading in that the ICER results for memantine are combined with those for the AChEI's, and this may not be valid. However, a crude plot such as that below raises the suspicion that there is a structural floor in the cost-utility achievable within the model. This needs to be addressed, since a disease

modifying treatment characterised by an increasing effect size over time must necessarily increase the time before institutionalisation. Figure 2



What is of particular concern to the present reviewer is that even if a treatment were developed which achieved what would generally be considered to be the highly desirable objective of a 50% reduction in rate of disease progression over 12 months (eg Vellas et al., *Disease-modifying trials in Alzheimer's disease: a European task force consensus*. Lancet Neurol 2007; 6:56–62; Report of the Lewin Group to the US Alzheimer's Association, 2007), this may still prove not to be cost effective according to the terms of reference of the present report. Unfortunately, the PenTAG model as it stands does not permit this question to be answered, as it does not provide a basis for distinguishing between the present symptomatic treatments and disease modifying treatments which can be expected to emerge in near future.

03 August 2010