Predicting the outcome of cholinesterase inhibitor treatment in Alzheimer’s disease

P J Connelly, N P Prentice, K G Fowler

See Editorial Commentary, p 305


Objective: To investigate the possibility that response to cholinesterase inhibitor therapy could be predicted by easily measurable variables that are known to change as a result of treatment (such as the Mini Mental State Examination), measures of function (such as the instrumental activities of daily living and the social behaviour subscales of the Nurse’s Observation Scale for Geriatric Patients), and measures of attention (such as the Digit Symbol Substitution Test; DSST), or that might influence response through structural (for example, age, cerebrovascular disease, medial temporal lobe (MTL) atrophy, hypertension) or chemical (for example, smoking) mechanisms.

Method: This was a cohort study of 160 consecutive outpatients with probable Alzheimer’s disease who commenced cholinesterase inhibitor treatment over a 3 year period in a semi-rural area of Scotland.

Results: The overall response rate was 42.1%. Stratification of response between good and poor responders was possible using baseline DSST and a measure of MTL thickness using CT. Among the patients, 60.4% of those above the cut off point for both DSST and MTL thickness (29/48 subjects) were classified as good responders, compared with 6.3% of subjects below the cut off point for both (1/16 subjects). Subjects above the cut off point for both measures were more likely to be classified as good responders than subjects with only one or no values above the respective cut off points ($\chi^2 = 10.61, df = 1, p = 0.001$)

Conclusions: The DSST and a measure of MTL thickness derived from CT scanning may be useful in improving the prediction of response to cholinesterase inhibitors in subjects with AD. Subjects with low DSST scores and more severe MTL atrophy are unlikely to respond to treatment. These preliminary data justify a prospective trial of the usefulness of our suggested predictive measures.

Variables that might predict the outcome of cholinesterase inhibitors include measurements shown to change in double blind randomised placebo controlled trials, including cognition,1–4 levels of functioning,5–7 and behaviour.3 Parameters that might affect response through possible structural mechanisms, (such as age, level of hippocampal atrophy and vascular change), or chemical mechanisms (such as smoking), could also help to discriminate responders and non-responders.

Baseline measures, such as severity of cognitive impairment,4 rate of pre-treatment decline,3 older age,3 current smoking,7 and the presence of concurrent vascular risk factors,8 have been reported as influencing response in separate studies. The role of the E4 allele of apolipoprotein E is controversial. Impaired response,9 no effect,10 and improved response11 have all been associated with its presence. Hippocampal volume was associated with response to a single dose of tacrine in one study,12 but not in a larger study with stricter outcome criteria by the same group.13 An inverse relationship between atrophy of the substantia innominata and response to 6 months’ treatment with donepezil has been described,14 although only the Mini Mental State Examination (MMSE)15 was used as an outcome measure.

Improvement in attentional rather than memory measures following chronic tacrine treatment has been reported,16 but the authors did not relate measures of baseline performance to ultimate outcome. Improvement in attention after single doses of tacrine17 18 did not correlate with response over a 52 week period.18 An association between reduction in theta power following cholinesterase inhibitor administration and cognitive outcome has been demonstrated in two small studies.19 20

Cholinesterase inhibitors are expensive and have significant side effects. Patients, clinicians, and managers would benefit if these drugs could be targeted at those most likely to respond in a clinical setting. Our service in Perth and Kinross is in a large semi-rural area in central Scotland with no specialised memory clinic. The service has a local CT scanner but very restricted access to MRI or EEG.

In this cohort study of subjects >45 years with NINCDS-ADRDA probable Alzheimer’s disease (AD),21 who were treated with cholinesterase inhibitor treatment during a 3 year period, we wished to establish whether any baseline variable or combination of variables could discriminate eventual “good” and “poor” responders after 6 months of treatment, using inexpensive, rapid, and effective tests, including neuroimaging. Additionally, while recognising the possible aetiological role of vascular changes in AD, we wished to examine the effect of vascular risk factors, such as hypertension and smoking, and the presence of vascular change on CT scan.

MATERIALS AND METHODS

Consecutive subjects with probable AD who scored 11–26 on the MMSE and <2 on the Rosen scale,22 with no evidence of significant cerebrovascular disease (see below) were commenced on either donepezil, rivastigmine, or galantamine. There were no restrictions on the choice of drug. Subjects with focal neurological (long tract) signs and symptoms were

Abbreviations: AD, Alzheimer’s disease; DSST, Digit Symbol Substitution Test; IADL, instrumental activities of daily living; MMSE, Mini Mental State Examination; MoM, multiple of the median; MTL, medial temporal lobe; NOSGER, Nurse’s Observation Scale for Geriatric Patients; S8, social behaviour
assessed by the Nurse’s Observation Scale for Geriatric
related value, using the formula of Jobst
then converted to the multiple of the median (MoM) age
to the medial wall of the temporal lobe, avoiding the choroid
sured the hippocampus directly from the hippocampal fissure
angle to the orbitomeatal line, or 15–20 degrees reverse angle
limitations, this equates to scanning 25–30 degrees reverse
Jobst
callipers. Standard axial views were augmented by temporal
thereafter. Measures of intra-cerebral structures were per-
scanner prior to February 2001 and a GE Hispeed NXI
CT scans were undertaken using a Toshiba XvisionGX
Kew cognitive map, 27 Digit Symbol Substitution Test
(dsst), 28 verbal fluency, finger tap, bimanual reciprocal co-
ordinat and motor sequence. 29 Functional ability was
assessed by the Nurse’s Observation Scale for Geriatric
Patiens (NOSGER), 30 completed by the carer. Each item on
the subscales was scored 1 (“never”) to 5 (“all the time”),
except when the reverse was indicated by the scale’s authors.

Vascular change on CT was assessed by considering
distribution and severity of peri-ventricular white matter
change, including infarcts. We defined significant cerebro-
vascular disease as vascular change on CT in subjects with
marked gait disturbance, frequent falls, urinary frequency or
urgency, 31 and poor performance of Luria motor tests 24–26 and
tests of frontal lobe function. These subjects were excluded
from treatment. Subjects with radiological vascular changes
in the absence of these symptoms and signs were included.

At baseline, cognition was assessed using the MMSE, 16
Kew cognitive map, 27 Digit Symbol Substitution Test
(DSST), 28 verbal fluency, finger tap, bimanual reciprocal co-
ordinat and motor sequence. 29 Functional ability was
assessed by the Nurse’s Observation Scale for Geriatric
Patiens (NOSGER), 30 completed by the carer. Each item on
the subscales was scored 1 (“never”) to 5 (“all the time”),
except when the reverse was indicated by the scale’s authors.

Medial temporal lobe (MTL) measurement
CT scans were undertaken using a Toshiba XvisionGX
scanner prior to February 2001 and a GE Hispeed NXI
thereafter. Measures of intra-cerebral structures were performed
on screen during analysis of the scan using electronic
callipers. Standard axial views were augmented by temporal
lobe orientated slices similar to the technique described by
Jobst et al. 11 Allowing for patient position and technical
limitations, this equates to scanning 25–30 degrees reverse angle
to the orbitomeatal line, or 15–20 degrees reverse angle
to the anthropological line, which approximates to a line
from the inner tip of the nasal cartilage to the pituitary fossa.
As recommended, 2 mm slices were performed in this
orientation, avoiding the orbits. “Minimal thickness” measured the hippocampus directly from the hippocampal fissure to the medial wall of the temporal lobe, avoiding the choroid plexus. The lesser of the left or right minimum values was then converted to the multiple of the median (MoM) age related value, using the formula of Jobst et al., 11 to control for the influence of age on the absolute value of the measure. This is a well validated measure that we felt would allow quantitative assessment of atrophy, and be more accurate than semi-qualitative measures in determining the influence of atrophy on response to treatment. We termed this variable MoM1.

From the same series of slices, a further measurement of the full thickness of the MTL was made across the parahippocampal gyrus and white matter arising directly from the hippocampus, which avoided sulcal and partial volume effects—that is, a measurement of the medial part of the temporal lobe from the medial wall of the lateral ventricle to the ambient cistern, avoiding the choroid plexus. Empirically, we converted the value of the average of the right and left thickness obtained by this measurement to a multiple of the median using the same formula and values for age related norms as in the calculation of MoM1. We believe that taking an average rating reduces adverse influences such as rotation of the subject’s head during the scanning procedure. This variable was termed MoM2. Although a learning curve exists for radiologists and radiographers, adequate views and consistency of measure-
ment can be obtained after about 20–25 scans.

All measurements were undertaken by a single radiologist
experienced in the procedure, who was blinded to the
subject’s diagnosis and severity of dementia. Intra-rater
reliability of the MoM2 measurement was assessed by re-
evaluation of 25 scans—that is, 50 measurements. Intra-class
correlation was 0.8 (p < 0.0005).

Dose titration
Standard titration protocols were used during the first
3 months. All subjects were being treated with donepezil
5 mg daily, rivastigmine 6 mg daily, or galantamine 16 mg
daily at this point. Subjects who had improved continued
their current dose. Those who had deteriorated were with-
drawn, in keeping with clinical practice. Those who had not
improved had the dose increased to 10 mg/day donepezil, or
the highest tolerated dose of rivastigmine (usually 4.5 mg
twice daily) or galantamine (usually 12 mg twice daily).

Outcome measures
Response was categorised after 6 months of treatment. To
minimise practice effects on test performance, subjects were
not tested in the latter 3 months of this period. Subjects were
assessed at the same time of day on each visit. At each follow
up, the MMSE and DSST were undertaken by researchers
blinded to the results of the carer’s assessment on the
NOSGER. Scores were not fed back to the patient or carer
until global outcome had been agreed.

Positive cognitive change was defined as a gain of at least
two points over the baseline MMSE score. 32 Positive
behavioural change was defined as an improvement of the
combined score on the instrumental activities of daily living
(IADL) and social behaviour (SB) subscales of the NOSGER,
or maintenance of maximum score. Positive global change
was defined as a tripartite agreement amongst doctor,
subject, and carer that appreciable improvement had
occurred, based upon global impressions. The total number
of variables rated as positive (“positive outcome criteria”) were calculated for each subject. “Good response” was
defined as the presence of at least two out of three positive
outcome criteria.

Statistical analysis
Age, MoM1, MoM2, MMSE, DSST, IADL, and SB scores all
had normal distribution (Kolmogorov-Smirnov test). Paired
T test analyses assessed differences between MMSE, DSST,
IADL, and SB scores at baseline and 6 months, in the cohort
as a whole and within the subgroups of “good” or “poor”
responders. Independent T test statistics were used to
cmpare these subgroups with each other.

Logistic regression procedure
Baseline MMSE, DSST, IADL, and SB scores were compared
with outcome classification using receiver operational char-
acteristic analysis. Each variable was dichotomised at the
point that maximised sensitivity and specificity. Age was
dichotomised into groups > 75 v ≤ 75 years. MoM1 values
were dichotomised using the value 0.79, the suggested cut off
point for diagnosing AD, 33 and MoM2 dichotomised around
the age related mean for MoM (1.00). CT appearances of
vascular change were dichotomised into “present” or
“absent”. Non-smokers were defined as those not using
cigars or cigarettes for at least 6 months, although in fact, all
ex-smokers had not smoked for a much longer period.
Variables had their predictive usefulness assessed on the
cohort who completed 6 months’ treatment. Significant odds
ratios (OR) are reported. All analyses were undertaken using
SPSS for Windows (version 10.0).
RESULTS
In the study, 160 subjects commenced treatment, but seven could not tolerate treatment for more than 1 week and were not entered into the cohort database. Of the remaining 153, six were withdrawn within 3 months because of intolerance; five were withdrawn after 3 months because of deterioration in their mental state; one was withdrawn after 3 months because of non-compliance, and one left the catchment area and was lost to follow up. Analysis was thus performed on 140 subjects completing 6 months’ treatment, 126 of whom had MoM2 measurement performed.

Mean (SD) age of the cohort (48 men, 92 women) was 76.30 (7.97) years at baseline; 17.1% were current smokers and 30% had either current hypertension or were receiving anti-hypertensive treatment. Baseline mean (SD) scores were: MMSE (22.58 (3.53), IADL (16.88 (4.37), SB (17.99 (4.78), and DSST (16.58 (11.70). Mean (SD) scores at 6 month follow up were: MMSE 22.80 (4.63), IADL 16.59 (4.62), SB (17.98 (5.15), and DSST 16.30 (12.46). Although none of the changes was significant at 6 months, there was a significant improvement over baseline in MMSE (p = 0.0005) and IADL (p = 0.025) after 3 months. The 13 subjects who either did not complete 3 months’ treatment or who were withdrawn after 3 months’ treatment did not differ significantly on any baseline measure from the cohort completing 6 months’ treatment.

Outcome
Using our criteria, 59 subjects (42.1%) of the cohort showed good response after 6 months’ treatment (38.6% of the 153 subjects tolerating treatment for more than 1 week). Of the 81 subjects (57.9%) who showed poor response, 17 had lower scores on each of MMSE, IADL, and SB scales than at baseline. By definition, differences in cognition and function between good and poor responders following treatment would be expected. Table 1 illustrates the differences between the groups.

Logistic regression analysis
Only MoM2 (OR 3.16; 95% confidence interval (CI) 1.17 to 8.51) and DSST (2.85; 1.42 to 5.73) showed significant odds ratios in the completed cases. Increased age, current smoking, hypertension, and vascular change on CT did not generate significant OR.

Improving prediction
DSST scores correctly predicted more response categories than MoM2 values in the logistic regression analysis. However, these variables had a different pattern of sensitivity and specificity. A DSST score ≥18 had a sensitivity for good response of 0.59 (95% CI 0.44 to 0.73) and a specificity of 0.66 (95% CI 0.53 to 0.77). MoM2 values ≥ the age adjusted median had a sensitivity for good response of 0.89 (95% CI 0.75 to 0.96) and a specificity of 0.30 (95% CI 0.18 to 0.43). Correlation between the MoM2 and DSST scores was only r = 0.072 (n = 126; p = 0.1 Pearson’s r, two tailed). Response was therefore tabulated initially using the cut off point for DSST scores, then by cut off point for MoM2. Results are shown in fig 1. Scores above the cut off point for both variables were associated with a higher likelihood of good response than if one or both were below the cut off point (χ² = 10.61, df = 1, p = 0.001). Although the absolute increase in prediction of good response increased only to 60.4% from 42.1% by undertaking the DSST and MoM2, the difference in the proportion of good responders with high DSST and low atrophy (29/48) and those with one or both below the cut point (26/77) was sufficient for the study to

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Comparison of good and poor outcome groups after 6 months’ treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline Good</td>
</tr>
<tr>
<td>MMSE</td>
<td>22.56 (3.36) (59)</td>
</tr>
<tr>
<td>IADL</td>
<td>16.42 (4.69) (59)</td>
</tr>
<tr>
<td>SB</td>
<td>18.53 (4.65) (59)</td>
</tr>
<tr>
<td>DSST</td>
<td>18.96 (12.14) (59)</td>
</tr>
<tr>
<td>MoM1</td>
<td>0.67 (0.16) (48)</td>
</tr>
<tr>
<td>MoM2</td>
<td>1.17 (0.19) (55)</td>
</tr>
<tr>
<td>AGE</td>
<td>75.84 (8.37) (59)</td>
</tr>
</tbody>
</table>

Values are mean (SD) (number). Two tailed p values are reported throughout.

Figure 1  Effect of dichotomising DSST score and MoM2 scores on outcome in subjects completing 6 months treatment. Some subjects did not have MTL thickness measured.
have greater than 80% power at this sample size (two tailed \( \alpha = 0.05 \)).

**DISCUSSION**

In this study, a combination of baseline performance on the DSST and the degree of medial temporal lobe atrophy appeared to distinguish subjects who had a good response to cholinesterase inhibitors from those with poor response. These variables may differ from those that predict responders to treatment with an active drug from placebo response. The poor response group showed decline in cognition that was similar to untreated patients.\(^{35} \; 36\) Decline in function of untreated patients does not appear to have been described. Therefore, we cannot comment on the rate of decline in poor responders relative to untreated patients. A 20% decline in function over 1 year has been described in placebo treated subjects.\(^{35} \; 36\) Our data demonstrate the differences in cognition and function between good and poor response that might be expected in a clinical setting. These differences are recognisable to patients and carers, and are clinically significant.

The DSST measures the speed of processing of information,\(^{37}\) which is in part dependent upon cholinergic transmission.\(^{38}\) Processing speed could therefore be expected to improve in those who respond to cholinesterase inhibitors, and it is encouraging that the DSST has some predictive utility in determining response. Although DSST scores have relatively large standard deviations at any given age range,\(^{39}\) the test is rapid and easy to perform. We recommend further assessment of this measure, and exploration of the effect of processing speed on response.

Our measure of MTL atrophy was obtained using an adaptation of CT scanning, which is widely accessible by psychiatrists dealing with the elderly in the UK and is relatively inexpensive. Lesser MTL atrophy was associated with better response.

Although the accuracy of MTL measurement as described by Jobst et al\(^{31}\) has been questioned,\(^{39}\) our measure (MoM2) is not of minimum thickness and may be easier to reproduce in other centres than MoM1. Given the variability in reports of atrophy of MTL in AD,\(^{40} \; 41\) we support the view that standardisation of MTL measurement is required.

The MMSE, although widely used in drug studies, is not designed for that purpose. This study shows that the MMSE at baseline is not a good predictor of response. However, the number of patients whose MMSE improved by at least one point after 6 months (53.7%) is substantially greater than that reported in a longitudinal study of untreated patients,\(^{42}\) suggesting that the improvement is unlikely to be due to the natural history of AD.

Potential risk factors such as older age, hypertension, and current smoking did not generate a significant OR in our study. The number of current smokers was low, but similar to the Scottish prevalence in this age group. We did not assess the influence of vascular risk on response,\(^{43}\) but instead assessed significance of vascular change using a combination of neuroradiographic appearance and clinical features. Although we recognise that CT scanning is less sensitive than MRI in detecting vascular lesions,\(^{44}\) lesions seen on CT have proved more accurate at predicting the development of neurological signs\(^{45}\) and are associated with neurological symptoms.\(^{46}\) The OR associated with vascular change that is not considered clinically significant using our definition would suggest that vascular change on CT scan per se is not associated with an altered rate of response.

Using a very restrictive cut off point for vascular risk factors on the Rosen scale\(^{47}\) and excluding subjects meeting our definition of clinically significant vascular disease, we believe that our population represents AD that is not coloured by vascular pathology. It is disappointing, therefore, that response rates to cholinesterase inhibitors were not better than we found, even though dropout rates in our clinical setting were much lower than expected based on published trials. The characteristics of our population, pattern of MMSE response,\(^{48} \; 49\) and response rate are similar to those reported in clinical trials\(^{10} \; 13\) and UK memory clinics.\(^{11} \; 12\) Therefore, it may be possible to generalise the results of this study to a wider population.

Ideally, the usefulness of our measure of MTL thickness requires assessment in other centres, and a fully prospective trial comparing the actual outcome of patients with none, one, or two of our suggested predictive factors using raters blinded to baseline classification would be justified by our study. Such a trial would reduce the problems associated with the analysis of multiple variables. Until then, given the current sensitivity and specificity of our measures, we would advise caution in using the DSST and MoM2 as the only measures to determine whether treatment is indicated for an individual patient. The combination of the DSST and MoM2 could be argued to have a relatively small effect compared to the use of no measures. However, a response rate of 60% is similar to that for antipsychotics or antidepressants, and may encourage prescription of cholinesterase inhibitors. We recognise that our measures appear to be a better predictor of poor than good response. Because the number of subjects with scores below the cut off point of both the DSST and MoM2 is small, caution must be used in restricting treatment to this group, despite only one of our subjects in this category showing good response at 6 months.

We have not assessed the usefulness of our predictive items in subjects with other types of dementia and cannot recommend their use as predictors of response to cholinesterase inhibitors in the absence of a clear diagnosis of AD. We hope that our study, in a clinical setting, stimulates the seeking of clinical or radiological parameters that might be more predictive of good response. The identification of these parameters is now vital if the benefits of cholinesterase inhibitors are to be maximised.

**ACKNOWLEDGEMENTS**

We acknowledge the contribution of Professor I C Reid, Department of Psychiatry, University of Aberdeen, who provided critical review of previous drafts of this paper and supervised the research work. Dr R C Chen, Senior Lecturer, Department of Epidemiology and Public Health, University of Dundee provided statistical advice. Our colleague Dr G Cousland contributed to the data collection. Invaluable secretarial support was provided by N Neil.

**Authors’ affiliations**

P J Connelly, Department of Psychiatry, University of Dundee, Murray Royal Hospital, Muirhall Road, Perth, PH2 7BH, UK

N P Prentice, Senior Lecturer in Old Age Psychiatry, Department of Psychiatry, University of Dundee, Murray Royal Hospital, Muirhall Road, Perth, PH2 7BH, UK

K G Fowler, Consultant Radiologist, Perth Royal Infirmary, Taymount Terrace, Perth PH1 1NX, UK

There was no funding source or sponsor for this work.

Competing interests: P J Connelly and N P Prentice have received honoraria and travel sponsorship from manufacturers of currently licensed cholinesterase inhibitors. P J Connelly and N P Prentice have received research funds from Eisai/Pfizer. K G Fowler has no competing interests.

**REFERENCES**


Predicting the outcome of cholinesterase inhibitor treatment in Alzheimer's disease

P J Connelly, N P Prentice and K G Fowler

J Neurol Neurosurg Psychiatry 2005 76: 320-324
doi: 10.1136/jnnp.2004.043539

Updated information and services can be found at:
http://jnnp.bmj.com/content/76/3/320

These include:

References
This article cites 36 articles, 10 of which you can access for free at:
http://jnnp.bmj.com/content/76/3/320#BIBL

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections

Drugs: CNS (not psychiatric) (1945)
Hypertension (380)
Radiology (1747)
Radiology (diagnostics) (1309)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/