

## REVIEW

# Size of the treatment effect on cognition of cholinesterase inhibition in Alzheimer's disease

K Rockwood

*J Neurol Neurosurg Psychiatry* 2004;**75**:677–685. doi: 10.1136/jnnp.2003.029074

**Background:** Six cholinesterase inhibitors (ChEIs) have been tested in people with Alzheimer's disease, using methods currently required for regulatory approval. The clinical importance of their treatment effects is controversial.

**Objective:** To determine whether cholinesterase inhibition produces treatment effects in Alzheimer's disease that are large enough to be clinically detectable.

**Methods:** Overview analysis of published trials of ChEIs in which the Alzheimer's Disease Assessment Scale—Cognitive Subscale (ADAS-Cog) and a global clinical measure were primary outcomes. Two quantitative summary measures of the treatment effect (Cohen's *d* and the standardised response mean (SRM)) were calculated and presented as funnel plots. Observed cases analyses and intention to treat (ITT) with the last observation carried forward (LOCF) analyses were compared.

**Results:** The median Cohen's *d* effect sizes (ES) using ITT samples with LOCF for the ADAS-Cog were: low dose of a ChEI (*n*=8 studies) median ES=0.15, range=0.03–0.22; medium dose (*n*=13) median ES=0.23, range=0.12–0.29; high dose (*n*=9) median ES=0.28, range=0.01–0.31. In general, the ES were larger when calculated as SRMs (for example, high dose ChEI studies, median SRM=0.47; range=0.30–0.63) and highest in the observed cases analyses (for example, high dose median SRM=0.56, range=0.35–0.78).

Global clinical scales produced similar estimates of ES (for example, high dose ChEI, ITT/LOCF median Cohen's *d*=0.29, range=0.20–0.47).

**Conclusions:** ChEIs produce small-moderate effect sizes in clinical trials which are reproducible and demonstrate a dose response. Better descriptions of the patterns of treatment response are needed to guide individual patient decisions about the effectiveness of treatment, but group effects are evident and appear large enough to be clinically detectable.

Correspondence to:  
K Rockwood, Centre for  
Health Care of the Elderly,  
5955 Veterans' Memorial  
Lane, Suite 1421, Halifax,  
Nova Scotia, Canada B3H  
2E1; kenneth.rockwood@  
dal.ca

Received 2 October 2003  
In revised form  
5 December 2003  
Accepted 21 January 2004

Although cholinesterase inhibitors (ChEIs) have been used for several years to treat Alzheimer's disease (AD), scepticism about their effectiveness persists. Many trials are short, the outcome measures are problematic, trials patients are generally healthier than most with the disease, and the treatment effects seem small.<sup>1–12</sup> Moreover, there is concern about how the data have been analysed.<sup>13</sup> Given that decline can be expected even in short trials, and that ChEIs engender more side effects than placebos, intention to treat (ITT) analyses that employ a last observation carried forward (LOCF) method of handling missing data are suspect. Briefly, the concern is about so-called dropout bias,<sup>13</sup> which occurs when patients who withdraw have systematically different outcomes than those who do not. While it is easy to imagine circumstances under which an LOCF analysis would be conservative—for example, if patients actively treated withdraw before they can recover from their illness, in the setting of AD trials, dropout bias could actually favour treatment. This would be the case if patients on active treatment more often withdrew early, due to side effects and before they had a chance to decline, whereas patients on placebo withdrew late and more often because of inefficacy. In consequence, LOCF analyses could actually compare patients on treatment with better scores with patients on placebo with worse scores.

Concern has also been expressed about the varying mixes of mildly and moderately impaired patients in the trials.<sup>12</sup> As moderately impaired patients decline more quickly than do mildly impaired ones,<sup>14 15</sup> any treatment effect is likely to be more demonstrable with that group. Indeed, it has been demonstrated that the rate of progression of dementia

predicts the response to cholinesterase inhibition.<sup>16</sup> Thus, from the standpoint of interpreting clinical meaningfulness, using trial data based largely on patients with moderately severe dementia to counsel patients with mild dementia is problematic.

Attention also has been drawn to the difficulty in generalising to daily clinical practice data from clinical trials carried out for regulatory purposes.<sup>1–3</sup> The outcome measures used to evaluate ChEIs are not much used in clinical practice. The primary outcomes usually have been the Alzheimer's Disease Assessment Scale—Cognitive Subscale (ADAS-Cog)<sup>17</sup> and a global clinical measure, typically the Clinician's Interview-Based Impression of Change Plus Caregiver Input (CIBIC-Plus). Tests such as the ADAS-Cog and CIBIC-Plus are seen as surrogate measures compared with "harder" measures, such as a delay in institutionalisation. The Cochrane Collaboration review of donepezil, for example, concluded that "the practical importance of these changes [in the ADAS-Cog] to patients and carers is unclear". Even though the CIBIC-Plus scores "clearly show some impact on global clinical states ..." "the rating scale is crude and difficult to interpret".<sup>1</sup>

This paper seeks to better understand the clinical meaningfulness of ChEIs in AD from the published data. It follows

**Abbreviations:** AD, Alzheimer's disease; ADAS-Cog, Alzheimer's Disease Assessment Scale—Cognitive Subscale; ChEIs, cholinesterase inhibitors; CIBIC-Plus, Clinician's Interview-Based Impression of Change Plus Caregiver Input; ITT, intention to treat; LOCF, last observation carried forward; OC, observed cases; SRM, standardised response mean

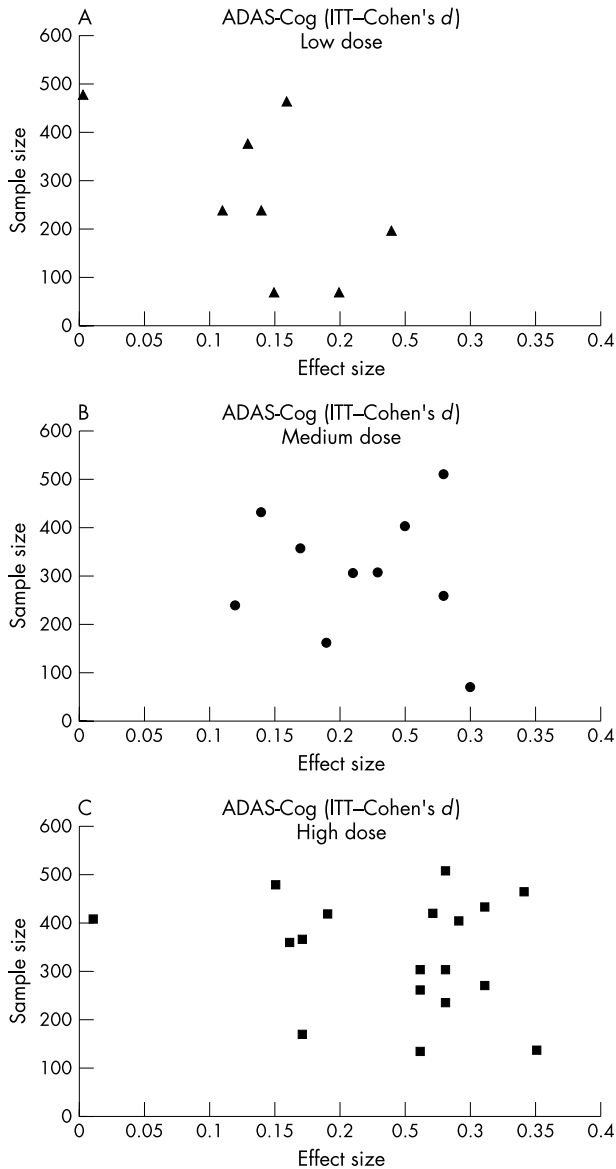
**Table 1** Overview of randomised controlled trials of cholinesterase inhibitors which used the ADAS-Cog and a global clinical measure

Trial	Proportion screened to enrolled	Number studied	Number withdrew	Duration	Secondary measures	Comments
<b>Tacrine</b> Davis <i>et al</i> , 1992 <sup>26</sup>	Not stated; 632 enrolled in "enrichment" phase	112 placebo	401 unable to calculate a "best dose"	6 week titration phase	MMSE, PDS, PSMS, IADL, ADAS-Noncog, ADAS-total	Included a pretest "enrichment phase" (n=632) to determine whether each patient had a "best dose" response (n=231) Global measure: CGIC
Knapp <i>et al</i> , 1994 <sup>27</sup>	Not stated	103 "best dose" (40 mg or 80 mg) 184 placebo	16 did not enter the double blind phase 59 placebo	2 week placebo baseline phase 6 weeks double blind phase 6 weeks open label phase 24 week titration phase	MMSE, PDS, GDS, Caregiver CGIC, IADL, FCCA, ADAS-total, ADAS-Noncog, PSMS	Global measure: CIBI
<b>Meirifonate</b> Becker <i>et al</i> , 1996 <sup>26</sup>	53.51	61 80 mg/d 179 120 mg/d 239 160 mg/d 23 placebo	32 80 mg/d 121 120 mg/d 172 160 mg/d 1 meirifonate	6 week double blind phase 3 week placebo run-in	MMSE, ADIC, ADAS-cog, ADAS-total	Included a 3 week loading dose followed by a 9 week maintenance dose Global measure: GIS
Cummings <i>et al</i> , 1998 <sup>27</sup>	Not stated	27 2.1 mg/kg weekly dose 120 placebo		3 month double blind phase Final open label phase 2 week screening 12 week double blind phase	PSMS, MMSE, IADL, CIBIS+, GERRI	Included a 2 week loading dose followed by a 10 week maintenance dose Global measure: CIBIC+
Morris <i>et al</i> , 1998 <sup>28</sup>	511:408	121 0.2 mg/kg/d 120 0.3 mg/kg/d 119 0.65 mg/kg/d 135 placebo	16 placebo	40 week optional open label phase 2 week screening	MMSE, NPI, GDS, ADAS-Noncog, CIBIS+, DAD	Included a 2 week loading dose followed by a 24 week maintenance dose Global measure: CIBIC+
<b>Donepezil</b> Rogers <i>et al</i> , 1996 <sup>21</sup>	Not stated	273 0.65 mg/kg/d 40 placebo	58 0.65 mg/kg/d 5 placebo	26 week double blind phase 8 week post treatment 12 week double blind 2 week placebo washout phase	MMSE, ADL, CDR-SB, QoL-P, QoL-C	Global measure: CGIC
Rogers <i>et al</i> , 1998a <sup>23</sup>	Not stated	42 1 mg/d 40 3 mg/d 39 5 mg/d 162 placebo 154 5 mg/d 157 10 mg/d	8 1 mg/d 2 3 mg/d 5 5 mg/d N/A	24 week double blind phase 6 week placebo washout phase	MMSE, CDR-SB, GoL	Global measure: CIBIC+
Rogers <i>et al</i> , 1998b <sup>22</sup>	Not stated	153 placebo 157 5 mg/d 158 10 mg/d 134 placebo	N/A	12 week double blind phase 3 week single blind washout	MMSE, CDR-SB, GoL	Global measures: CIBIC+
Homma <i>et al</i> , 2000 <sup>24</sup>	Not stated	134 5 mg/d 274 placebo 271 5 mg/d 273 10 mg/d	39 withdrawals N/A	24 week double blind study 24 week double blind phase 6 week placebo washout phase	CDR-SB, CMCS, MENFIS IDDD, CDR-SB, GoL	ADAS-J Cog—Japanese version of the ADAS-Cog Global measure: J-CGIC Global measure: CIBIC+
Burns <i>et al</i> , 1999 <sup>25</sup>	Not stated					

Table 1 (continued)

Trial	Proportion screened to enrolled	Number studied	Number withdrew	Duration	Secondary measures	Comments
<b>Rivastigmine</b> Rösler <i>et al</i> , 1999 <sup>30</sup>	831:725	239 placebo 243 1-4 mg/d 243 6-12 mg/d 235 placebo	31 placebo 33 1-4 mg/d 79 6-12 mg/d 38 placebo	26 weeks	PDS, GDS, MMSE	Global measure: CIBIC+
Corey-Bloom <i>et al</i> , 1998 <sup>29</sup>	925:699	233 placebo 233 1-4 mg/d 231 6-12 mg/d	34 1-4 mg/d 82 6-12 mg/d	7 week titration phase 19 week flexible dose phase	PDS, MMSE, GDS	Global measure: CIBIC+
<b>Galantamine</b> Tariot <i>et al</i> , 2000 <sup>31</sup>	1178:978	286 placebo 140 8 mg/d 279 16 mg/d 273 24 mg/d 213 placebo 212 24 mg/d 211 32 mg/d 125 placebo 261 24-32 mg/d 215 placebo 220 24 mg/d 218 32 mg/d	46 placebo 32 8 mg/d 60 16 mg/d 61 24 mg/d 41 placebo 68 24 mg/d 89 32 mg/d 12 placebo 86 24-32 mg/d 29 placebo 44 24 mg/d 55 32 mg/d	4 week placebo run-in phase 5 month double blind phase 4 week placebo run-in phase 6 month double blind phase 4 week placebo run-in period 3 month double blind phase 4 week placebo run-in period 5 month double blind phase 2 week placebo run-in 2 week dose escalation 10 weeks on assigned dose	ADCS/ADL, NPI	Global measure: CIBIC+
Raskind <i>et al</i> , 2000 <sup>32</sup>	764:636	213 placebo 212 24 mg/d 211 32 mg/d 125 placebo 261 24-32 mg/d 215 placebo 220 24 mg/d 218 32 mg/d	41 placebo 68 24 mg/d 89 32 mg/d 12 placebo 86 24-32 mg/d 29 placebo 44 24 mg/d 55 32 mg/d	4 week placebo run-in phase 6 month double blind phase	ADL, DAD	Global measure: CIBIC+
Rockwood <i>et al</i> , 2001 <sup>34</sup>	534:386	125 placebo 261 24-32 mg/d 215 placebo 220 24 mg/d 218 32 mg/d	12 placebo 86 24-32 mg/d 29 placebo 44 24 mg/d 55 32 mg/d	4 week placebo run-in period 3 month double blind phase 4 week placebo run-in period 5 month double blind phase	NPI, ADAS-Noncog, DAD, PSQI	Global measure: CIBIC+
Wilcock <i>et al</i> , 2000 <sup>33</sup>	753:653	215 placebo 220 24 mg/d 218 32 mg/d	29 placebo 44 24 mg/d 55 32 mg/d	4 week placebo run-in period 5 month double blind phase	DAD, 13-item ADAS-Cog, MMSE	Global measure: CIBIC+
Wilkinson <i>et al</i> 2001 <sup>35</sup>	Not stated	88 18 mg/d 56 24 mg/d 54 36 mg/d	14 placebo 25 18 mg/d 14 24 mg/d 26 36 mg/d	2 week placebo run-in 2 week dose escalation 10 weeks on assigned dose	PDS-1	Global measure: CGIC
<b>Physostigmine</b> Van Dyck <i>et al</i> , 2000 <sup>40</sup>	Not stated	93 placebo 83 best dose	11 placebo 17 best dose	3 week dose enrichment phase 4 week placebo washout phase 12 week double blind phase 4 week dose titration phase 2 week all placebo phase 6 week double blind phase 2 week washout period 9 week dose titration phase 24 week double blind phase	MMSE, IADL, GCIC	Global measure: CIBIC+
Thal <i>et al</i> , 1996 <sup>39</sup>	1111:366	183 placebo 183 best dose	N/A	4 week dose titration phase 2 week all placebo phase 6 week double blind phase 2 week washout period 9 week dose titration phase 24 week double blind phase	MMSE, PSMS, IADL	Global measure: CGIC
Thal <i>et al</i> , 1999 <sup>38</sup>	699:475	117 placebo 176 30 mg/d 182 36 mg/d	31 placebo 110 30 mg/d 124 36 mg/d	2 week washout period 9 week dose titration phase 24 week double blind phase	IADL, GERRI, CGIC	Global measure: CIBIC+

ADAS, Alzheimer's Disease Assessment Scale; ADCS/ADL, Alzheimer's Disease Cooperative Study Activities of Daily Living Inventory; ADL, Activities of Daily Living checklist; CDR-SB, Clinical Dementia Rating Scale—Sum of the Boxes; CGIC, Clinical Global Impression of Change; CIBI, Clinician Interview-Based Impression of Change Plus Caregiver Input; CMCS, Caregiver-rated Modified Crichton Scale; d, day; DAD, Disability Assessment for Dementia; FCCA, Final Comprehensive Consensus Assessment; GDS, Global Deterioration Scale; GERRI, Geriatric Evaluation of Relative's Rating Instrument; GIS, Global Improvement Scale; IADL, Instrumental Activities of Daily Living; IDDD, Interview for Deterioration of Daily activities in Dementia; MENDIS, Mental Function Impairment Scale; MMSE, Mine-Mental State Examination; N/A, not available; NPI, Neuropsychiatric Inventory; PDS, Progressive Deterioration Scale; PSMS, Physical Self-Maintenance Scale; PSQI, Pittsburg Sleep Quality Index; Gal-P, -, Quality of Life (Patient or Caregiver)



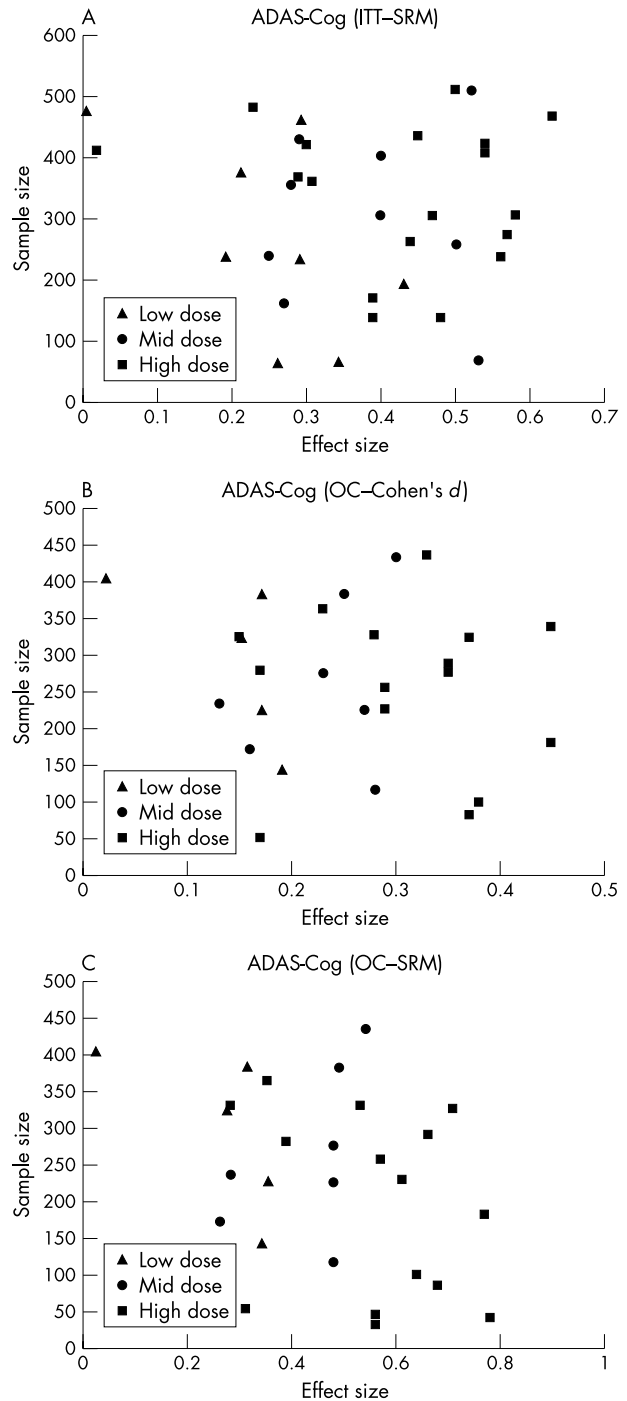
**Figure 1** ADAS-Cog estimates of the size of the treatment effect (Cohen's *d*, intention to treat) by dose (see text for definitions).

the suggestion that effects will be clinically meaningful if they are statistically significant, large enough to be detectable, demonstrate a dose-response effect, are reproducible, and evidently clinically observable.<sup>18</sup> We have also used the opportunity of systematically reviewing the data to gain some insight into the impact of LOCF analyses by comparing the effect sizes of ITT/LOCF with observed case analyses.

**METHODS**

**Search strategy**

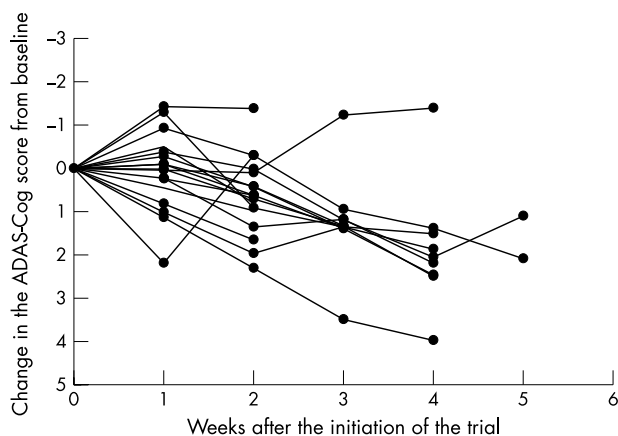
For the selection of published trials the search strategy of the Canadian Coordinating Office for Health Technology Assessment (CCOHTA) was followed.<sup>6</sup> Briefly, the CCOHTA selected trials by standard methodological criteria, which were then subjected to a blinded quality assessment using the Jadad Scale.<sup>19</sup> Only those with a quality score of 5 or more were accepted. To maximise comparability, only those ChEI trials that used the ADAS-Cog and a separate global clinical measure were included, and these were restricted to trials with community dwelling patients with mild moderate AD. Qizilbash has urged caution in meta-analyses in dementia.<sup>20</sup>



**Figure 2** Other ADAS-Cog estimates of the size of the treatment effect by type of effect size, and type of analysis, by low, mid (medium), and high doses.

Rather than a meta-analysis, a narrative review was undertaken with attention to trial selection, comparability and method of analysis, and with pre-specified criteria of the assessment of clinical meaningfulness.

Herein are summarised data for five trials of donepezil,<sup>21-25</sup> three of metrifonate<sup>26-28</sup> and two of rivastigmine.<sup>29-30</sup> Since the CCOHTA review was published, five additional trials of galantamine have become available and have been included.<sup>31-35</sup> We have also included the two tacrine trials that met all criteria<sup>36-37</sup> as did three trials of physostigmine.<sup>38-40</sup>



**Figure 3** Inter-trial variability on ADAS-Cog placebo responses in Alzheimer's disease trials.

**Operationalisation of the criteria to assess meaningfulness**

Statistical significance was assessed through an overview of the published literature, and held to be the case whenever the primary analysis demonstrated that the results were unlikely ( $p < 0.05$ ) to have been observed if the null hypothesis (of no difference between the active treatment and placebo groups) were true.

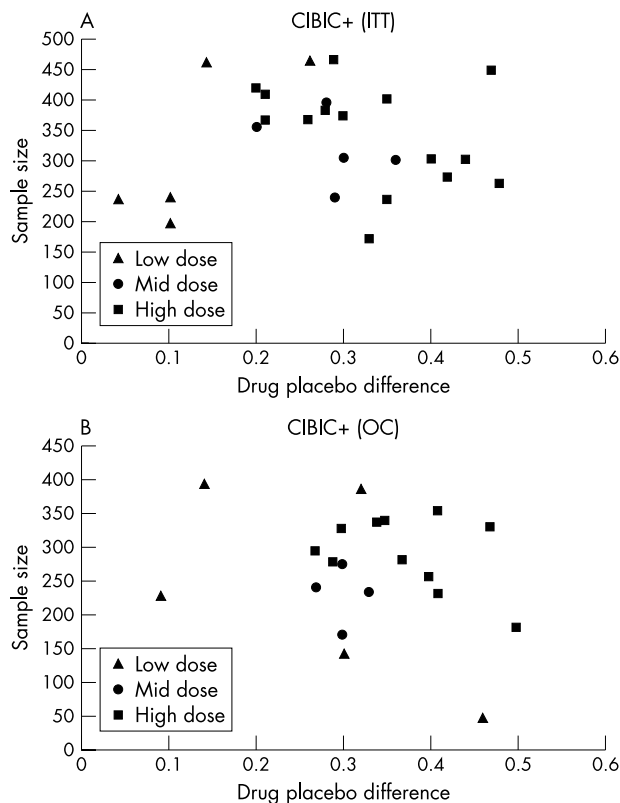
The criterion of being *large enough to be clinically detectable* was assessed by estimating effect sizes. Although the term is sometimes used simply to refer to the absolute differences between placebo and treatment groups,<sup>12</sup> a standard effect size relates this number to the observed variance in the

measures.<sup>41</sup> Effect sizes were calculated as either (or both, where available) Cohen's *d* or standardised response means (SRMs). Both Cohen's *d* and SRMs calculate a "signal", for any measure, as the mean difference at the endpoint between those on active treatment and those on placebo, given baseline differences (that is,  $x_{t_2-t_1} - y_{t_2-t_1}$ , where  $x$  is the active treatment group,  $y$  is the placebo group,  $t_2$  is the endpoint and  $t_1$  is the baseline). Both compare this signal in the average difference to an estimate of "noise": either the pooled baseline deviation (Cohen's *d*) or the pooled standard deviation of the change score (SRM). In the calculation of both Cohen's *d* and SRM, and given that higher scores of each of the CIBIC-Plus and ADAS-Cog suggest worse disease, effect sizes that favour treatment will have negative values. Each study showed favourable effects, so that, for convenience only the absolute values are presented.

On theoretical grounds, and given that the best bias is conservative—that is, least likely to show a treatment benefit, the choice between Cohen's *d* and SRM depends on assumptions about the characteristics of the sample under study. If the sample is artificially homogeneous at the outset (for example, as a result of strict inclusion/exclusion criteria) then the baseline variance will be artificially small, Cohen's *d* will be speciously large, and thus SRM would be preferred. On the other hand, if the sample becomes more homogeneous over time (for example, through survivor effects) the opposite would be the case and Cohen's *d* would be preferred. As this proposition has not been formally tested in anti-dementia trials, where possible, we present both estimates. Also presented are the absolute differences in relation to overall patterns of improvement or decline. We also estimated these measures with those observed to have completed the trial (observed cases; OC) and with the LOCF in an ITT analysis. Not every study precisely specifies the definition, but here the ITT sample is taken to comprise all patients who received at least one dose of double blind study medication, and for whom data are available from the baseline evaluation and at least one post-baseline assessment. An ITT/OC analysis includes only those patients who received at least one dose of study medication, had a baseline evaluation, and had a final evaluation at the scheduled end of the trial while on double blind study medication. This is also referred to as a "completers" analysis. An ITT/LOCF analysis consists of all patients who received at least one dose of study medication, and had at least one pretreatment evaluation and one evaluation after the start of treatment. The last observation is carried forward and used as a so-called endpoint evaluation.

Few published studies provide baseline standard deviation scores or standard deviations of the change scores, but these can be calculated from the sample size and standard errors of the mean treatment difference or confidence intervals. Whether these differences were large enough to be clinically detectable was considered on two grounds. Effect sizes greater than 0.20 are conventionally held to be clinically detectable.<sup>41</sup> In addition, effect sizes for individual trials were correlated with scores on the global clinical measure used in the trial.

Where a dose-response relationship holds, the effect size will show peaks for each dose. In the case of most ChEIs, two doses were assessed: a low dose and a medium-high dose. Where each higher dose is reported separately, separate effect sizes are presented, giving small, medium and high dose effect size estimates. The doses were as follows: donepezil, no low dose, medium dose 5 mg daily, high dose 10 mg daily; rivastigmine, low dose 1.5 mg twice daily, medium-high dose 3.0 mg twice daily or higher; metrifonate, low dose 0.2 mg/kg bodyweight, medium dose 0.3 mg/kg, high dose 0.65 mg/kg; galantamine, low dose 4 mg twice daily, medium dose 8 mg



**Figure 4** Size of the treatment difference in the CIBIC-PLUS raw score, by method of analysis.

twice daily, high dose 12 mg twice daily or higher. Note that to achieve the highest dose, there is interval dose titration.

The *demonstration of a dose response* was tested by comparing effect sizes of a given drug across doses of that drug, and likewise between drugs. These data are presented as funnel plots. Funnel plots display the effect size against the sample size.<sup>42</sup> They derive their name from the probability property that a sample estimate will more closely approximate the true value of the population parameter the larger the sample size, other things being equal. In consequence, the effect size will “funnel in” on the true value as the sample size increases.

*Reproducibility* was assessed first by examining the range of effect sizes between trials on the funnel plots. In addition, the distribution of the effect sizes across different ChEIs was observed. In each case, reproducibility was defined as comparability of the effects, considering variable sample size and trial duration, as portrayed in the funnel plots.

The criterion of being “evidently clinically observable” was operationalised as a statistically significant difference on the CIBIC-Plus or like global clinical measure. To summarise these measures, where no baseline variance is possible, and where change score variance terms are rarely presented, the effect size was calculated simply as the absolute difference in the proportions deteriorating, improving, or demonstrating no change. Their size and potential to show dose–response relationship can also be considered as measures of the reproducibility of the data.

## RESULTS

Table 1 presents an overview of the dosing and size of each trial, the ratio of patients screened to those enrolled and the number of withdrawals from each arm of the trial. While withdrawals vary between the trials, there are more withdrawals in the treatment arms than in the placebo arms. Figure 1 presents a funnel plot of the effect sizes of ChEIs on the ADAS-Cog for the ITT/LOCF using Cohen’s *d*, for the low, medium, and high doses (panels A–C, respectively). Statistical significance was obtained in each of the studies with an effect size of more than 0.15 and a sample size of more than 250 ( $p < 0.05$ ). With four exceptions (from one study each of tacrine, rivastigmine, galantamine, and physostigmine) the high doses of each study showed effect sizes of greater than 0.20 using Cohen’s *d* and the LOCF/ITT analysis. In every case, the effect size was greater than 0.20 with the highest dose using the SRM. Figure 2 is the funnel plot for the ITT/LOCF using the SRM (panel A) and for the OC (panels B, C) analyses. Individual data points are plotted for each drug at each dose. In the ITT/LOCF analyses using Cohen’s *d* as the estimate of effect size, peaks are detectable at about 0.15 (panel A) and 0.25, corresponding roughly to low and medium-high doses, respectively. The SRMs (fig 2, panel A) likewise show peaks, at about the median values of 0.26 (low dose) and 0.47 (high dose). As would be expected, the highest estimates are seen in the OC analyses (fig 2, panels B, C), with, for example, a median SRM in the OC medium-high dose studies being 0.56; (range 0.35 to 0.78).

Figure 3 provides some context for these estimates by noting the course of the ADAS-Cog changes in the placebo arms of the trials. With two exceptions, the trials show persistent declines in the placebo arms over time, although the rate of decline varies.

Figure 4 is a funnel plot of the effect sizes of ChEIs on the CIBIC-Plus for the ITT/LOCF (panel A) and OC (panel B) analyses. Individual data points are plotted for each drug at each dose. Again, two peaks are detectable, one about 0.15 and another about 0.25, corresponding to low and medium-high doses, respectively. In general, the global clinical measures produced estimates of effect sizes (where these could be calculated) that were similar to the ADAS-Cog—for

example, high dose ChEI, ITT/LOCF median SRM is 0.29 (range 0.20 to 0.47). The correlation of the ADAS-Cog effect sizes to the CIBIC+ effect sizes for ITT analyses was 0.29 for the ITT/LOCF analyses and 0.60 (SRM) for the OC analyses.

## DISCUSSION

This study is part of a general strategy to assay the clinical importance of ChEIs. The studies reviewed suggest that anti-dementia drug trials show statistically significant differences that are large enough to be clinically detectable, demonstrate a dose response, are reproducible, and thus have evident clinical meaningfulness.<sup>18</sup> Although controversy exists,<sup>43</sup> ChEIs appear to have a biologically plausible basis.<sup>44–45</sup> In general, ChEIs produce clinically detectable effects on groups of patients treated for AD. The effects appear to be broadly reproducible, to demonstrate a dose response, and to be large enough to be clinically detectable.

These conclusions must be made with important caveats. Not all the data are available from all trials, as some employed different measures of cognition. Still, the ADAS-Cog is now accepted as a standard, and decisions are being made about ChEIs on the basis of this measure, so it is useful to have some insight into its merit in drug trials. Another potential problem is publication bias. As far as can be determined, this report includes all known pivotal studies that employed the ADAS-Cog, for each of the three ChEIs widely used for AD, but we note that there are no estimates with small sample sizes and small effect sizes, which is often a sign of publication bias.<sup>42</sup>

Another limitation is the generalisability of clinical trial data. Necessarily, the only patients considered in this overview are those who were eligible to take part in clinical trials. Given that dementia and multiple medical comorbidities are all strongly age related,<sup>46</sup> many patients who suffer from dementia are not eligible for trials. In addition, many who meet an initial prescreen evaluation are excluded, so that only about 13% of patients referred for trials are enrolled.<sup>8</sup> “Real world” effectiveness is therefore a challenge in interpreting clinical meaningfulness,<sup>6</sup> although this is not limited to anti-dementia trials.<sup>47–50</sup>

The present analysis also helps provide some context for new studies. For example, data from two placebo controlled trials of donepezil have been published, which offer some insight into the size of the treatment effects for more than the maximum six months of the studies presented. Mohs *et al* studied 431 patients in a 54 week trial of donepezil versus placebo.<sup>51</sup> The study did not use the ADAS-Cog or CIBIC-Plus, but rather employed a survival analysis of decline in one or more basic Activities of Daily Living. The Mini-Mental State Examination (MMSE)<sup>52</sup> was the main cognitive measure. Cohen’s *d* was 0.13 using the OC analysis, and 0.20 using the LOCF analysis. Winblad *et al* used the MMSE in a 52 week placebo controlled trial of donepezil.<sup>53</sup> We calculated Cohen’s *d* to be 0.42 using the OC analysis and 0.57 using the LOCF analysis. Thus estimates of long term effects vary between these trials, which have important design differences. Differences in drug effects also appear to exist between domains. For example, a meta-analysis of the impact of ChEIs on behaviour showed that while most produced statistically significant differences, effect sizes typically were less than 0.20.<sup>54</sup>

Perhaps the most important caveat is that these group effects do not assist in the interpretation of a given effect in a given patient. The trials conducted for regulatory purposes are helpful in telling us that *something* is going on, but they do not tell us enough about what that is to provide much of a guide for individual patient decisions. This dilemma is perhaps best expressed by the dominance of the “no change” category on the CIBIC-Plus, which remains the single most

common summary outcome of treatment. It is generally not the case that patients exhibit no clinically detectable change. Rather, it appears to be that these patients experience changes that the treating clinicians judge to be offsetting—for example, cognitive improvement but functional decline, or functional improvement but behavioural worsening.<sup>55–56</sup> Perhaps buried in these summary scores are detectable patterns of change that may consist of reproducible effects of treatments. To mine these patterns, additional strategies are required.<sup>57–58</sup> Also, it appears that changes of a given magnitude in a measure such as the ADAS-Cog are variable—that is, a four point change indicating differing treatment effects depends on not just the severity of dementia, but also on the rate of disease progression.<sup>16</sup>

Still, analysis of effect sizes offers a means of understanding clinical meaningfulness by suggesting where it is likely to occur, and how likely it is to be evident to observers with varying degrees of expertise. It also allows comparison of the responsiveness (sensitivity to change) between measures within a given trial,<sup>59</sup> and provides some insight into the effects of new drugs or new indications. For example, the six month data<sup>60</sup> on the Ginkgo biloba extract EGb761 in the treatment of AD show effect sizes (Cohen's *d*) comparable to the effect sizes of cholinesterase inhibition (ADAS-Cog ITT 0.42 in OC analysis, and 0.33 with ITT/LOCF; Clinical Global Impression of Change (CGIC) 0.29 with OC analysis and 0.18 with ITT). In contrast, the 12 month data<sup>61</sup> show an effect size of 0.20 with the ADAS-Cog, and 0 with the CGIC. Similarly, in a six month study in moderate to severe AD, patients on donepezil showed a small degree of improvement on the CIBIC-Plus, compared with continued decline in the placebo control group, resulting in an effect size of 0.49 (ITT/LOCF; the OC analysis was 0.37).<sup>62</sup> The effect size (SRM) detected by the CIBIC-Plus in a memantine trial was 0.27 (the absolute difference was 0.3).<sup>63</sup> In a placebo controlled study of donepezil in vascular dementia, the effect sizes demonstrated by the ADAS-Cog ranged between 0.30 and 0.37, depending on the dose (5 mg or 10 mg) or whether the LOCF or endpoint analysis was used.<sup>64</sup> These estimates were similar to those seen in another trial using galantamine, in which the six month treatment effect size estimated by the ADAS-Cog was 0.35.<sup>65</sup>

The effect size method thus allows for some quantification, and takes into account response in the placebo group, as does the estimate of the “number needed to treat”<sup>66</sup> which also can be seen as a variant of an effect size calculation for proportions.<sup>42</sup> These methods thus seem to be superior to estimation of the “minimal clinically important difference” which estimates signal better than noise, and which does not readily capture large placebo related changes.<sup>67</sup>

As noted, anti-dementia trials face the particular challenge of likely deterioration with or without treatment, and thus the problem of dropout bias, especially when early discontinuation is due to side effects related to medication. If patients exposed to drugs withdraw due to side effects, the effectiveness of the drug would be overestimated by an analysis which did not take dropout into account. In general, the remedy for this is an ITT analysis, employing either of two methods to handle missing data. In the ITT-non-completer equals failure method any patient whose data are missing is considered a treatment failure. Perhaps it was the experience of toxicity with the early ChEIs that persuaded the Cochrane Collaboration reviewers of ChEIs to be sceptical about the otherwise supposed, conservative bias (with respect to placebo effects) of an ITT/LOCF analyses. Although, in theory, early withdrawal can serve either to give a smaller or a larger effect than would have been the case had patients stayed in the trial, the present overview showed the ITT/LOCF analyses to be more conservative than the ITT/OC case

analyses. A preferred method is the retrieved dropout analyses, in which patient measures are taken at the scheduled endpoint of the trial, whether they stayed in the study or not. This is not yet the standard for anti-dementia trials.

This study also helps contribute to our understanding of how to interpret the CIBIC-Plus. Although for the grouped data the effect size estimates of the CIBIC-Plus and ADAS-Cog overlapped, the correlations between them were only notional for the ITT/LOCF analyses (~0.30) but higher for the OC analyses (0.60). This suggests both that each measure assays different aspects of the treatment effect,<sup>58</sup> in keeping with the original rationale for the use of a global clinical measure,<sup>68–69</sup> and that patients who successfully complete studies might have recognisable profiles. The latter is important as clinicians (and thus the CIBIC-Plus) appear to be better at capturing deterioration than at capturing improvement,<sup>70</sup> especially if improvement in some domains is seen in the face of deterioration in others.<sup>55–56</sup> Such considerations point to a larger and compelling problem, which is not well addressed by the current studies, of how we get to grips with the idea of treatment success in AD, even when that success falls short of cure.<sup>71</sup> As argued elsewhere, one important step in getting to grips with how to evaluate this aspect of clinical meaningfulness would be to assay the preferences of patients with AD, and the preferences of their caregivers.<sup>57–58–71</sup>

Estimation of the achievement of patient and caregiver preferences in anti dementia trials might also allow a better means of comparing drug treatment effects with effects that arise from non-pharmacological interventions. For example, in a meta-analysis of psychosocial interventions for caregivers of people with dementia, Brodaty and colleagues showed that such interventions yielded caregiver benefits of about 0.30 for any main outcome.<sup>72</sup> Types of outcomes varied, being most successful in helping with the patient's mood, and least successful in helping with caregivers' feelings of burden. Comparing like outcomes with like is important if we are to adjudicate the concern that drug treatment for dementia is claiming more than its fair share of resources.<sup>73</sup>

In short, with reasonable consistency, ChEIs produce small-moderate effect sizes in clinical trials. A dose response can be demonstrated, and the patterns appear to be replicable both within and across trials. Dropout bias remains a concern in understanding group differences. Better descriptions of the patterns of treatment response are needed if we are to be guided in individual patient decisions about the effectiveness of treatment.

## ACKNOWLEDGEMENTS

The author is grateful to Heather Merry, MSc, for her assistance in preparing the funnel plots and in calculating effect sizes from the published data.

KR is supported by an Investigator award from the Canadian Institutes of Health Research and by the Dalhousie Medical Research Foundation as the Kathryn Allen Weldon Professor of Alzheimer Research.

Competing interests: The author has conducted trials and/or received research and/or consulting and/or special fees from each of the following companies whose products are included in the review: Bayer; Janssen-Ortho; Novartis; Pfizer.

## REFERENCES

- 1 Birks JS, Melzer D, Beppu H. Donepezil for mild and moderate Alzheimer's disease. *Cochrane Database Syst Rev* 2000;(4):CD001190.
- 2 Birks J, Grimley E, Lakovidou V, et al. Rivastigmine for Alzheimer's disease (Cochrane Review). In: *Cochrane Library*, Issue 2. Oxford: Update Software, 2001.
- 3 Bentham P, Gray R, Hills R, et al. Twelve-week response to cholinesterase inhibitors does not predict future benefit: the AD2000 experience [abstract]. *Neurobiol Aging* 2002;(suppl 1):337.

- 4 Olin J, Schneider L. Galantamine for Alzheimer's disease. *Cochrane Database Syst Rev* 2002;(3):CD001747.
- 5 Clegg A, Bryant J, Nicholson T, et al. Clinical and cost-effectiveness of donepezil, rivastigmine, and galantamine for Alzheimer's disease. A systematic review. *Int J Technol Assess Health Care* 2002;18:497-507.
- 6 Husereau D, Wolfson C, Shukla VK. Drug treatments for Alzheimer's disease: efficacy, outcome measurements and cost-effectiveness. *Technology Overview no 4*. Canadian Coordinating Office for Health Technology Assessment, 2001.
- 7 Frisoni GB. Treatment of Alzheimer's disease with acetylcholinesterase inhibitors: bridging the gap between evidence and practice. *J Neurol* 2001;248:551-7.
- 8 Treves TA, Verchovsky R, Klimovitsky S, et al. Recruitment rate to drug trials for dementia of the Alzheimer type. *Alzheimer Dis Assoc Disord* 2000;14:209-11.
- 9 Qizilbash N, Whitehead A, Higgins J, et al. Cholinesterase inhibition for Alzheimer disease: a meta-analysis of the tacrine trials. Dementia Trialists' Collaboration. *JAMA* 1998;280:1777-82.
- 10 Pryse-Phillips W. Do we have drugs for dementia? *Arch Neurol* 1999;56:735-7.
- 11 Mayeux R, Sano M. Treatment of Alzheimer's disease. *N Engl J Med* 1999;341:1670-9.
- 12 McLendon BM, Doraiswamy PM. Defining meaningful change in Alzheimer's disease trials: the donepezil experience. *J Geriatr Psychiatry Neurol* 1999;12:39-48.
- 13 Hills R, Gray R, Stowe R. Drop-out bias undermines findings of improved functionality with cholinesterase inhibitors [abstract]. *Neurobiol Aging* 2002;23(suppl 1):338.
- 14 Jacobs D, Sano M, Marcker K, et al. Age at onset of Alzheimer's disease: relation to pattern of cognitive dysfunction and rate of decline. *Neurology* 1994;44:1215-20.
- 15 Mitnitski A, Graham JE, Mogilner AJ, et al. The rate of decline in function in Alzheimer's disease and other dementias. *J Gerontol Med Sci* 1999;54:M65-M69.
- 16 Farlow MR, Hake A, Messina J, et al. Response of patients with Alzheimer disease to rivastigmine treatment is predicted by the rate of disease progression. *Arch Neurol* 2001;58:417-22.
- 17 Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. *Am J Psychiatr* 1984;141:1356-64.
- 18 Rockwood K, MacKnight C. Assessing the clinical importance of statistically significant improvement in anti-dementia drug trials. *Neuroepidemiology* 2001;20:51-6.
- 19 Oremus M, Wolfson C, Perrault A, et al. Interrater reliability of the modified Jada quality scale for systematic reviews of Alzheimer's disease drug trials. *Dement Geriatr Cogn Disord* 2001;12:232-6.
- 20 Qizilbash N. Systematic reviews. In: Qizilbash N, Schneider LS, Chui H, eds. *Evidence-based Dementia Practice*. Oxford: Blackwell Publishing, 2002.
- 21 Rogers SL, Friedhoff LT, the Donepezil Study Group. The efficacy and safety of donepezil in patients with Alzheimer's disease: results of a US multicentre, randomized, double-blind, placebo-controlled trial. *Dementia* 1996;7:293-303.
- 22 Rogers SL, Doody RS, Mohs RC, et al. Donepezil improves cognition and global function in Alzheimer disease. *Arch Intern Med* 1998;158:1021-31.
- 23 Rogers SL, Farlow MR, Doody RS, et al. A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Alzheimer's disease. *Neurology* 1998;50:136-45.
- 24 Homma A, Takeda M, Imai Y, et al. Clinical efficacy and safety of donepezil on cognitive and global function in patients with Alzheimer's disease. *Dement Geriatr Cogn Disord* 2000;11:299-313.
- 25 Burns A, Rossor M, Hecker J, et al. The effects of donepezil in Alzheimer's Disease—results from a multinational trial. *Dement Geriatr Cogn Disord* 1999;10:237-44.
- 26 Becker RE, Colliver JA, Markwell SJ, et al. Double-blind, placebo-controlled study of metrifonate, an acetylcholinesterase inhibitor, for Alzheimer disease. *Alzheimer Dis Assoc Disord* 1996;10:124-31.
- 27 Cummings JL, Cyrus PA, Bieber F, et al. Metrifonate treatment of the cognitive deficits of Alzheimer's disease. Metrifonate Study Group. *Neurology* 1998;50:1214-21.
- 28 Morris JC, Cyrus PA, Orazem J, et al. Metrifonate benefits cognitive, behavioral, and global function in patients with Alzheimer's disease. *Neurology* 1998;50:1222-30.
- 29 Corey-Bloom J, Anand R, Veach J, for the ENA 713 B352 Study Group. A randomized trial evaluating the efficacy and safety of ENA 713 (rivastigmine tartrate), a new acetylcholinesterase inhibitor, in patients with mild to moderately severe Alzheimer's disease. *Int J Geriatr Psychopharmacol* 1998;1:55-65.
- 30 Rösler M, Anand R, Cicin-Sain A, et al. Efficacy and safety of rivastigmine in patients with Alzheimer's disease: international randomised controlled trial. *BMJ* 1999;318:633-8.
- 31 Tariot PN, Solomon PR, Morris JC, et al. A 5-month, randomized, placebo-controlled trial of galantamine in AD. *Neurology* 2000;54:2269-76.
- 32 Raskind MA, Peskind ER, Wessel T, et al. A 5-month, randomized, placebo-controlled trial of galantamine in AD. The Galantamine USA-10 Study Group. *Neurology* 2000;54:2269-76.
- 33 Wilcock GK, Lilienfeld S, Gaens E, on behalf of the Galantamine International 1 Study Group. Efficacy and safety of galantamine in patients with mild to moderate Alzheimer's disease: multicentre randomised controlled trial. *BMJ* 2000;321:1-7.
- 34 Rockwood K, Mintzer J, Truyen L, et al. Effects of a flexible galantamine dose in Alzheimer's disease: a randomised, controlled trial. *J Neurol Neurosurg Psychiatry* 2001;71:589-95.
- 35 Wilkinson D, Murray J. Galantamine: a randomized, double-blind, dose comparison in patients with Alzheimer's disease. *Int J Geriatr Psychiatry* 2001;16:852-7.
- 36 Davis KL, Thal LJ, Gamzu ER, et al. A double-blind, placebo-controlled multicenter study of tacrine for Alzheimer's disease. *N Engl J Med* 1992;327:1253-9.
- 37 Knapp MJ, Knopman DS, Solomon PR, et al. A 30-week randomized controlled trial of high-dose tacrine in patients with Alzheimer's disease. *JAMA* 1994;271:985-91.
- 38 Thal LJ, Ferguson JM, Mintzer J, et al. A 24-week randomized trial of controlled-release physostigmine in patients with Alzheimer's disease. *Neurology* 1999;52:1146-52.
- 39 Thal LJ, Schwartz G, Sano M, et al. A multicenter double-blind study of controlled-release physostigmine for the treatment of symptoms secondary to Alzheimer's disease. Physostigmine Study Group. *Neurology* 1996;47:1389-95.
- 40 Van Dyck CH, Newhouse P, Falk WE, et al. Extended-release physostigmine in Alzheimer disease: a multicenter, double-blind, 12-week study with dose enrichment. Physostigmine Study Group. *Arch Gen Psychiatry* 2000;57:157-64.
- 41 Cohen J. *Statistical Power Analyses for the Behavioural Sciences*, 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.
- 42 Light RJ, Pillimer DB. *Summing Up: The science of reviewing research*. Cambridge, MA: Harvard University Press, 1994.
- 43 Davis KL, Mohs RC, Marin D, et al. Cholinergic markers in elderly patients with early signs of Alzheimer disease. *JAMA* 1999;281:1401-6.
- 44 Francis PT, Palmer AM, Snape M, et al. The cholinergic hypothesis of Alzheimer's disease: a review of progress. *J Neurol Neurosurg Psychiatry* 1999;66:137-47.
- 45 DeKosky ST, Ikonovic MD, Styren SD, et al. Upregulation of choline acetyltransferase activity in hippocampus and frontal cortex of elderly subjects with mild cognitive impairment. *Ann Neurol* 2002;51:145-55.
- 46 McDowell I. Alzheimer's disease: insights from epidemiology. *Aging (Milano)* 2001;13:143-62.
- 47 Bahit MC, Cannon CP, Antman EM, et al. Direct comparison of characteristics, treatment, and outcomes of patients enrolled versus patients not enrolled in a clinical trial at centers participating in the TIMI 9 Trial and TIMI 9 Registry. *Am Heart J* 2003;145:109-17.
- 48 Sokka T, Pinsky T. Eligibility of patients in routine care for major clinical trials of anti-tumor necrosis factor alpha agents in rheumatoid arthritis. *Arthritis Rheum* 2003;48:313-18.
- 49 Licht RW. Limits of the applicability and generalizability of drug trials in mania. *Bipolar Disord* 2002;4(suppl 1):66-8.
- 50 Swanson GM, Bailiar JC III. Selection and description of cancer clinical trials participants—science or happenstance? *Cancer* 2002;95:950-9.
- 51 Mohs RC, Doody RS, Morris JC, et al. A 1-year, placebo-controlled preservation of function survival study of donepezil in AD patients. *Neurology* 2001;57:481-8.
- 52 Folstein MF, Folstein S, McHugh PR. Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189-98.
- 53 Winblad B, Engedal K, Soininen H, et al. A 1-year, randomized, placebo-controlled study of donepezil in patients with mild to moderate AD. *Neurology* 2001;57:489-95.
- 54 Trinh NH, Hoblyn J, Mohanty S, et al. Efficacy of cholinesterase inhibitors in the treatment of neuropsychiatric symptoms and functional impairment in Alzheimer disease: a meta-analysis. *JAMA* 2003;289:210-16.
- 55 Joffres C, Bucks RS, Haworth J, et al. Patterns of clinically detectable treatment effects with galantamine: a qualitative analysis. *Dement Geriatr Cogn Disord* 2003;15:26-33.
- 56 Joffres C, Graham J, Rockwood K. Qualitative analysis of the clinician interview-based impression of change (Plus): methodological issues and implications for clinical research. *Int Psychogeriatr* 2000;12:403-13.
- 57 Rockwood K, Graham JE, Fay S. ACADIE Investigators. Goal setting and attainment in Alzheimer's disease patients treated with donepezil. *J Neurol Neurosurg Psychiatry* 2002;73:500-7.
- 58 Rockwood K, Joffres C. Halifax Consensus Conference on Understanding the Effects of Dementia Treatment. Improving clinical descriptions to understand the effects of dementia treatment: consensus recommendations. *Int J Geriatr Psychiatry* 2002;17:1006-11.
- 59 Rockwood K, P Stolee. Responsiveness of outcome measures in an anti-dementia drug trial. *Alzheimer Dis Assoc Disord* 2000;14:182-5.
- 60 Le Bars PL, Katz MM, Berman N, et al. A placebo-controlled, double-blind, randomized trial of an extract of Ginkgo biloba for dementia. North American EGB Study Group. *JAMA* 1997;278:1327-32.
- 61 LeBars PL, Kiesel M, Itil KZ. A 26-week analysis of a double-blind, placebo-controlled trial of the ginkgo biloba extract EGB 761 in dementia. *Dement Geriatr Cogn Disord* 2000;11:230-7.
- 62 Feldman H, Gauthier S, Hecker J, et al. A 24-week, randomized, double-blind study of donepezil in moderate to severe Alzheimer's disease. *Neurology* 2001;57:613-20.
- 63 Reisberg B, Doody R, Stoffler A, et al. Memantine in moderate-to-severe Alzheimer's disease. *N Engl J Med* 2003;348:1333-41.
- 64 Wilkinson D, Doody R, Helme R, et al. Donepezil in vascular dementia: A randomized, placebo-controlled study. *Neurology* 2003;61:479-86.
- 65 Erkinjuntti T, Kurz A, Gauthier S, et al. Efficacy of galantamine in probable vascular dementia and Alzheimer's disease combined with cerebrovascular disease: a randomised trial. *Lancet* 2002;359:1283-90.
- 66 Livingstone G, Katona C. How useful are cholinesterase inhibitors in the treatment of Alzheimer's disease? A number needed to treat analysis. *Int J Geriatr Psychiatry* 2000;15:203-7.



- 67 **Burback D**, Molnar FJ, St John P, *et al.* Key methodological features of randomized controlled trials of Alzheimer's disease therapy. *Dement Geriatr Cogn Disord* 1999;**10**:534–40.
- 68 **Schneider LS**, Olin JT, Doody RS, *et al.* Validity and reliability of the Alzheimer's Disease Cooperative Study-Clinical Global Impression of Change. *Alzheimer Dis Assoc Disord* 1997;**11**(suppl 2):S22–32.
- 69 **Knopman DS**, Knapp MJ, Gracon SJ, *et al.* The Clinician Interview-Based Impression (CIBI): a clinician's global change rating scale in Alzheimer's disease. *Neurology* 1994;**44**:2315–21.
- 70 **Quinn J**, Moore M, Benson DF, *et al.* A videotaped CIBIC for dementia patients: validity and reliability in a simulated clinical trial. *Neurology* 2002;**58**:433–7.
- 71 **Rockwood K**, Wallack M, Tallis R. The treatment of Alzheimer's disease: understanding success short of cure. *Lancet Neurol* 2003;**2**:630–3.
- 72 **Brodsky H**, Green A, Koschera A. Meta-analysis of psychosocial interventions for caregivers of people with dementia. *J Am Geriatr Soc* 2003;**51**:657–64.
- 73 **Heller T**, Heller L. First among equals? Does drug treatment claim more than its fair share of resources? *Dementia* 2003;**2**:7–19.

### Clinical Evidence—Call for contributors

*Clinical Evidence* is a regularly updated evidence based journal available worldwide both as a paper version and on the internet. *Clinical Evidence* needs to recruit a number of new contributors. Contributors are health care professionals or epidemiologists with experience in evidence based medicine and the ability to write in a concise and structured way.

#### Currently, we are interested in finding contributors with an interest in the following clinical areas:

Altitude sickness; Autism; Basal cell carcinoma; Breast feeding; Carbon monoxide poisoning; Cervical cancer; Cystic fibrosis; Ectopic pregnancy; Grief/bereavement; Halitosis; Hodgkins disease; Infectious mononucleosis (glandular fever); Kidney stones; Malignant melanoma (metastatic); Mesothelioma; Myeloma; Ovarian cyst; Pancreatitis (acute); Pancreatitis (chronic); Polymyalgia rheumatica; Post-partum haemorrhage; Pulmonary embolism; Recurrent miscarriage; Repetitive strain injury; Scoliosis; Seasonal affective disorder; Squint; Systemic lupus erythematosus; Testicular cancer; Varicocele; Viral meningitis; Vitiligo

However, we are always looking for others, so do not let this list discourage you.

#### Being a contributor involves:

- Appraising the results of literature searches (performed by our Information Specialists) to identify high quality evidence for inclusion in the journal.
- Writing to a highly structured template (about 2000–3000 words), using evidence from selected studies, within 6–8 weeks of receiving the literature search results.
- Working with *Clinical Evidence* Editors to ensure that the text meets rigorous epidemiological and style standards.
- Updating the text every eight months to incorporate new evidence.
- Expanding the topic to include new questions once every 12–18 months.

If you would like to become a contributor for *Clinical Evidence* or require more information about what this involves please send your contact details and a copy of your CV, clearly stating the clinical area you are interested in, to Claire Folkes (cfolkes@bmjgroup.com).

### Call for peer reviewers

*Clinical Evidence* also needs to recruit a number of new peer reviewers specifically with an interest in the clinical areas stated above, and also others related to general practice. Peer reviewers are health care professionals or epidemiologists with experience in evidence based medicine. As a peer reviewer you would be asked for your views on the clinical relevance, validity, and accessibility of specific topics within the journal, and their usefulness to the intended audience (international generalists and health care professionals, possibly with limited statistical knowledge). Topics are usually 2000–3000 words in length and we would ask you to review between 2–5 topics per year. The peer review process takes place throughout the year, and our turnaround time for each review is ideally 10–14 days.

If you are interested in becoming a peer reviewer for *Clinical Evidence*, please complete the peer review questionnaire at [www.clinicalevidence.com](http://www.clinicalevidence.com) or contact Claire Folkes (cfolkes@bmjgroup.com).