How to appraise an article on therapy

James P. Warner and Robert Blizard

This first of a series of articles intended to guide clinicians through the evidence-based process assesses the published evidence for drug therapy in Alzheimer's disease. A Medline search identified a randomised, controlled trial of donepezil. Although this trial appeared to have a reasonably sound methodology, we had some doubts about the treatment effect and applicability of the results.

In recent years there has been a move to using evidence-based practices in psychiatry (Goldner & Bilsker, 1995; Geddes & Harrison, 1997). An academic programme involving case conferences and journal clubs provides a platform for the introduction of evidence-based medicine (EBM) (Gilbody, 1996; Warner & King, 1997). Fundamental to the EBM process is the sequence (a) setting a question prompted by a clinical senario; (b) defining a literature search strategy and undertaking a search and (c) critically appraising the relevant paper(s) (Sackett *et al*, 1997).

What is the evidence for treating early Alzheimer's disease with donepezil?

Vignette

A 69-year-old retired civil servant presented with an 18 month history of gradual memory impairment and increasing apathy. There was no relevant past medical or psychiatric history. He did not appear depressed, and scored 21/30 on the Mini Mental State Examination (MMSE; Folstein *et al*, 1975). Subsequent neuropsychological testing showed a global reduction in cognitive function compat-ible with dementia. A diagnosis of probable Alzheimer's disease was made. His wife had read about a new drug for treating dementia called donepezil, and requested her husband be prescribed this.

Question

In individuals with Alzheimer's disease what is the evidence that donepezil improves cognitive function compared to standard treatment?

Literature search

We had a vague recollection of at least one randomised controlled trial being published, but could not recall any details. We therefore performed a literature search on Medline using the keywords 'Alzheimer's disease'. This identified 6610 references between 1993 and 1997. This is clearly too many references to sift through so the search was restricted to articles on 'therapy' (n=161) and then restricted further to 'randomised controlled trials'. Four randomised controlled trials on therapy in Alzheimer's disease were identified, but none were about donezepil. Trying a different tack we did a textword search on 'donepezil'. This yielded two references only, including a randomised controlled trial (Rogers *et al*, 1996): 'The safety and efficacy of donepezil in patients with Alzheimer's disease: results of a US multicentre, randomised, double-blind, placebo-controlled trial'.

Getting the article

This paper is not held in our library, but was obtained in three days through the library at a cost of £3.00.

Brief outline of the article

Rogers *et al* (1996) begin with an outline of the prevalence of Alzheimer's disease and rationale for using cholinesterase inhibitors in the treatment of this disorder. Subjects were selected if they were between 55 and 85, met robust diagnostic criteria for Alzheimer's and had a MMSE of between 10 and 26. They had to be fully ambulatory. There were rigorous exclusion criteria including: no other psychiatric/neurological disorder; no gastrointestinal, renal, hepatic or cardiovascular disease; and no history of alcohol or drug misuse.

Patients were randomised to one of four groups: donepezil 1 mg, 3 mg, 5 mg or placebo. They were given the drug for 12 weeks and assessed at baseline, 1,3,6,9,12 weeks and again after a fortnight washout at 14 weeks. Primary outcome measures were the cognitive sub-scale of the Alzheimer's Disease Assessment Scale (ADAS-cog; Rosen *et al*, 1984) and the Clinical

https://doi.org/10.1192/pb.22.6.342 Published online by Cambridge University Press

Global Impression of Change (CGIC; Guy, 1976). The MMSE was an ancillary outcome measure. ADAS-cog has a score range of 0 to 70, and the MMSE a score of 0 to 30. The CGIC is rated by a doctor, who subjectively assesses the patient at each visit and rates an impression of change compared to baseline.

Of the 161 patients randomised, 141 completed treatment. Differences in ADAS-cog scores between placebo and donepezil 5 mg groups at week 12 was 3.2 ($P<0.001$), favouring donepezil. The majority of patients in all groups did not deteriorate as measured by the CGIC. The difference in MMSE score between the placebo group and donepezil 5 mg group of 0.8 points was not statistically significant. Donepezil was well tolerated, the incidence of adverse effects being similar to placebo. The authors concluded that donepezil 5 mg daily for 12 weeks provided "significant clinical improvements in cognitive function".

Critical appraisal of an article on therapy

This followed the recommendations of Sackett *et al* (1997). There are three main elements: Is the trial valid? What do the results show? Can I apply the results to my patients?

Is the trial valid?

Was assignment of patients randomised? Yes. The authors state patients were randomised, but no details of method are given. Details, such as method of randomisation and blinding, although of considerable importance, are often omitted because of constraints of space.

Were all patients accounted for at the end of the trial? No. The authors document clearly the number and reasons for patient withdrawal but there are minor discrepancies between these numbers and the figures presented in the main outcome table. All 161 patients are included in the endpoint intention-to-treat analysis, but there were slight differences in the numbers in each group at endpoint, compared to the start of the trial.

Were patients and clinicians blind to the treatment? Yes. The authors state the study was double-blind.

Were the groups treated equally? Yes.

Were all groups similar at the start of the trial? No. There were minor differences in body height and weight between groups. This is acknowledged by the authors and is unlikely to contribute to the differences observed.

What are the results?

How large was the treatment effect? This is best addressed by calculating the number needed to treat (NNT) (i.e. how many patients need to be treated in order for one patient to derive significant benefit?) Using the results at the endpoint, the main outcome measure, the ADAS-cog shows an advantage for donepezil 5 mg compared with placebo of 3.2 units of change from baseline. Lower doses produce smaller improvements, consistent with a linear dose response. However, discrete outcome (e.g. better or not), which are necessary for calculating the NNT, are not provided on the cognitive measurements in this paper. A proxy measure for patient outcome is provided by the CGIC scores, where treatment failure is defined as a score of between 5 and 7. Using this measure 8/40 (20%) patients on placebo were classed as treatment failures, compared with 4/38 (11%) on donepezil 5 mg. It is helpful to construct a 2×2 table (see Table 1). The absolute risk reduction (ARR) is the difference in risk between the placebo and active group. In this study, the ARR for getting worse as measured by the CGIC is $0.2 - 0.11 = 0.09$. The NNT, the reciprocal of the ARR, is 11. In other words, on the basis of this study 11 patients need to be treated with donepezil rather than placebo for one patient to not deteriorate as measured by the CGIC.

How precise is the treatment effect? The authors do not apply confidence intervals on any outcome variables, and do not provide the standard errors necessary for the reader to calculate them. Figure 1 (of Rogers *et al*, 1996) provides standard error bars, but these are not displayed for all groups.

Can I apply this evidence to my patient?

Is my patient very different to those in the trial? No, but they could easily have been. The sample in this study were 94% Caucasian and were all very healthy (see exclusion criteria).

How great is the potential benefit for my patient? Do not know. This study does not

Table 1. 2×2 table of outcome of the clinician's global impression of change at endpoint (+, improved or unchanged; −, worse (treatment failure))

	Outcome	
	+	−
Placebo	32	8
Donepezil 5 mg	34	4

https://doi.org/10.1192/pb.22.6.342 Published online by Cambridge University Press

convince us that the benefits derived from donezepil outweigh the harms and costs. Our patient would expect a mean improvement of less than one point on the MMSE after 12 weeks treatment. It is unclear whether a 3.2 point improvement on the 70-point ADAS-cog is clinically important. According to this study, donepezil appears safe, but as the cost per patient is about £1000 per year, clinicians may consider the potential benefits are outweighed by the costs. Further information on long-term safety, efficacy and a cost–benefit analysis are needed to answer this question.

Comment

The first problem we highlight is the possibility that key papers may be missed on a literature search (Greenhalgh, 1997). Key elements in identifying appropriate articles include carefully defining the initial question and using an appropriate search strategy with correct headings. Eliciting the help of a medical librarian will often considerably improve your effectiveness when searching databases. The paper in question (Rogers *et al*, 1996) should have been identified by the initial search strategy, but was missed because it is not logged on Medline as a randomised controlled trial. If the searcher was unaware of the name of the drug in order to do a textword search, the article could have been missed. Other databases, such as Embase and Psychlit may well identify papers not appearing in Medline. For example, articles in the *Psychiatric Bulletin* are listed in Embase; a medline search would fail to find the article you are reading now.

At first sight, the article appraised here appears to have several flaws. However, it is not a definitive study and several factors, like the range of donepezil dose, wide span of cognitive impairment in the subjects and relatively short time of the study may mask more significant benefits of donepezil. A consequence of using the EBM approach is that results are likely to be treated with greater circumspection. Most authors will present the data in a way that suggests a positive outcome, for example a difference in scores between groups that is statistically very significant. The advantage of the critical appraisal outlined above is that the calculation of the NNT provides the clinician with a figure that is more meaningful than the data as presented in the article. It also helps to highlight issues of validity and the applicability of the study in question. On the basis of the paper appraised, we concluded that the evidence underpinning the use of donepezil is not compelling. Further, hitherto unpublished studies suggest that donepezil has greater utility in treating mild to moderate Alzheimer's disease than this initial study, with a number needed to treat nearer 5 than 11. This highlights one of the core tenets of EBM; that knowledge shifts rapidly and needs to be updated. Indeed, by the time this article is published there are likely to be more published studies on donepezil available in the literature.

Some published studies do not present data in such a way that allows calculation of the NNT, or appraisal of the validity and applicability of the study. Of those that do, many appear to be disappointing when appraised in this way. After a time a sense of nihilism can develop, and it is important to maintain a sense of perspective. Research is hard to do and it is easy to be critical with hindsight. The important messages that can be derived from the process are a sense of the quality of the paper and, for those engaged in research, ideas to improve the quality of future studies and the presentation of the results.

References

FOLSTEIN, M. F., FOLSTEIN, S. E. & MCHUGH, P. R. (1975) 'Mini Mental State.' A practical method for grading the cognitive state of subjects for the clinics. *Journal of Psychiatric Research*, **12**, 189–198.

GUY, W. (1976) *ECDEU Assessment Manual for Psychopharmacology*. Rockville, MD: NIMH.

GEDDES, J. R. & HARRISON, P. J. (1997) Closing the gap between research and practice. *British Journal of Psychiatry*, **171**, 220–225.

GILBODY, S. (1996) Evidence-based medicine. An improved format for journal clubs. *Psychiatric Bulletin*, **20**, 673–675.

GOLDNER, E. M. & BILSKER, D. (1995) Evidence-based psychiatry. *Canadian Journal of Psychiatry*, **40**, 97–101.

GREENHALGH, T. (1997) How to read a paper – the medline database. *British Medical Journal*, **315**, 180–183.

ROGERS, S. L., FRIEDHOFF, L. T. & THE DONEPEZIL STUDY GROUP (1996) The safety and efficacy of donepezil in patients with Alzheimer's disease: results of a US multicentre, randomised, double-blind, placebo-controlled trial. *Dementia*, **7**, 293–303.

ROSEN, W. G., MOHS, R. C. & DAVIS, K. L. (1984) A new rating scale for Alzheimer's disease. *American Journal of Psychiatry*, **141**, 1356–1364.

SACKETT, D. L., RICHARDSON, W. S., ROSENBERG, W., *et al* (1997) *Evidence-Based Medicine*. New York: Churchill Livingstone.

WARNER, J. P. & KING, M. B. (1997) Evidence-based medicine and the journal club: a cross-sectional survey of participants' views. *Psychiatric Bulletin*, **21**, 532–534.

*James Warner, *Lecturer, and* Robert Blizard, *Medical Statistician, Royal Free Hospital School of Medicine, Rowland Hill Street, London NW3 2QG*

*Correspondence

https://doi.org/10.1192/pb.22.6.342 Published online by Cambridge University Press