

Evaluating drug treatments for Parkinson's disease: how good are the trials?

Keith Wheatley, Rebecca L Stowe, Carl E Clarke, Robert K Hills, Adrian C Williams, Richard Gray

Keith Wheatley and colleagues make the case that most trials of drug treatment for Parkinson's disease have crucial methodological faults—and provide little reliable evidence on differences between classes of drugs

Birmingham
Clinical Trials Unit,
University of
Birmingham,
Birmingham
B15 2RR

Keith Wheatley
reader in medical
statistics

Rebecca L Stowe
information scientist

Robert K Hills
statistician

Richard Gray
professor of medical
statistics

City Hospital,
Birmingham
B18 7QH

Carl E Clarke
reader in clinical
neurology

Queen Elizabeth
Hospital,
Birmingham
B15 2TH

Adrian C Williams
professor of clinical
neurology

Correspondence to:
K Wheatley
k.wheatley@
bham.ac.uk

BMJ 2002;324:1508–11

Parkinson's disease is one of the commonest causes of disability in older people, with over 100 000 patients in the United Kingdom and at least 8000 new cases diagnosed annually. Prevalence and incidence will both increase with the ageing population and the reduction in competing causes of mortality such as stroke and coronary heart disease.¹ No cure currently exists, and medical treatment is directed towards alleviating symptoms.² Levodopa relieves symptoms in most patients with Parkinson's disease, but long term use of levodopa is associated with motor complications such as involuntary movements (dyskinesias), along with a shortened response to each dose (wearing-off phenomenon) and unpredictable "on-off" fluctuations. A number of other drugs have been used,³ either alone or with reduced doses of levodopa, in an attempt to delay the onset of motor complications in early Parkinson's disease or to control complications once they have developed. These agents have primarily been from three classes of drug: dopamine agonists, monoamine oxidase type B inhibitors, and catechol-O-methyltransferase inhibitors.

Many randomised controlled trials have evaluated these drugs, but uncertainty about their relative effectiveness remains. This review assesses the methods used in these trials to reveal the quality of the existing evidence base.

Identifying trials and extracting data

To identify publications from 1966 to the end of 2001 we searched the Cochrane Library, NHS Centre for Reviews and Dissemination, and Health Technology Assessment databases for systematic reviews and Medline, Embase, PubMed, and the Web of Science for primary research. We hand searched major journals in the field, including *Movement Disorders*, *Parkinsonism and Related Disorders*, *Neurology*, and *Journal of Neurology, Neurosurgery and Psychiatry*. We contacted experts in the field in attempt to identify studies not found through electronic and hand searching, and scanned reference lists of retrieved papers and websites relating to Parkinson's disease. No restriction on study design was made other than randomisation.

Summary points

The prevalence of Parkinson's disease will increase as the population ages, making it important to identify reliably the most effective drug therapy

Although many randomised controlled trials have evaluated the efficacy of different classes of drugs in both early and later Parkinson's disease, uncertainty about best treatment remains because of small numbers, inadequate follow up, and inappropriate end points

Much larger trials are needed with long term follow up and end points of relevance to patients

Large simple pragmatic trials have improved treatment of heart disease, stroke, and cancer and their methods should be applied to Parkinson's disease and other neurodegenerative diseases

Unpublished and non-randomised studies were not included.

Reporting of trials

The quality of reporting in many trials was poor. Often the randomisation procedure, the method of allocation concealment, the mechanism of blinding, and the number of patients included in analyses and numbers lost to follow up were inadequately described, making it difficult to exclude potential sources of bias. The results were often inconsistently reported compared with the analyses specified in the methods section and were often poorly described, with important statistical variables such as confidence intervals or significance values not given. Identifying multiple publications of the same trial was time consuming; many trials were published at several different stages and authors often failed to make this clear. The difficulty of ascertaining whether previously

Table 1 Numbers of trials and patients for comparisons of drug treatments in early and later Parkinson's disease

Comparison	No of trials*	No of patients	Mean No of patients per study	Mean duration of follow up (years)
Early Parkinson's disease:	43	8 177	190	3.8
Dopamine agonist v placebo (\pm levodopa)	13	1 407	108	1.0
Monoamine oxidase type B inhibitor v placebo (\pm levodopa)	11	1 445	131	3.4
Catechol- <i>O</i> -methyltransferase inhibitor v placebo (\pm levodopa)	2	381	191	0.8
Dopamine agonist (\pm levodopa) v levodopa	20	4 262	213	4.7
Monoamine oxidase type B inhibitor (\pm levodopa) v levodopa	2	852	426	7.2
Monoamine oxidase type B inhibitor v dopamine agonist (\pm levodopa)	2	385	192	2.6
Other†	6	1 041	174	6.4
Later Parkinson's disease‡:	67	6 454	96	0.5
Dopamine agonist v placebo	23	2 395	104	0.5
Monoamine oxidase type B inhibitor v placebo	18	540	30	0.5
Catechol- <i>O</i> -methyltransferase inhibitor v placebo	16	2 166	135	0.5
Dopamine agonist v dopamine agonist	9	1 348	150	0.5
Catechol- <i>O</i> -methyltransferase inhibitor v dopamine agonist	3	499	166	0.2
All trials*	110	14 631	133	2.3

*Five multiple comparison trials exist; in these, each patient counts towards more than one comparison but just once to the subtotals and total, which give the actual numbers of trials, numbers of patients randomised, and lengths of follow up.

†Dopamine agonist v dopamine agonist (\pm levodopa), dopamine agonist (+ monoamine oxidase type B inhibitor) v levodopa, dopamine agonist (+ monoamine oxidase type B inhibitor) v dopamine agonist, dopamine agonist (+ monoamine oxidase type B inhibitor) v monoamine oxidase type B inhibitor (+ levodopa).

‡All trials in later Parkinson's disease are on a background of levodopa based treatment.

published small pilot studies were included in the report of the main trial increases the likelihood of including patients twice in a meta-analysis. We hope that future trial reports will be of a higher standard and will adopt the CONSORT guidelines.⁴ Registration systems for notification of trials at their start would help to identify all randomised controlled trials and avoid publication bias.

Comparisons of trials

Table 1 outlines comparisons of drug treatments in early and later Parkinson's disease. Research in early Parkinson's disease has concentrated on comparing dopamine agonists with placebo and with levodopa and comparisons of monoamine oxidase type B inhibitors with placebo. In later Parkinson's disease, similar numbers of trials have compared each drug class with placebo but few trials have directly compared different classes of drug.

Numbers of trials and patients

Although almost 15 000 patients were randomised in the 110 published trials that were identified, the mean size was just 133 patients per trial (Table 1). The 43 trials in early Parkinson's disease had an average of 190 patients in each. Only four studies in early disease accrued more than 500 patients, and almost half of the trials included fewer than 100 patients (median 116). Trials of adjuvant therapy in later Parkinson's disease tended to be even smaller, with the 67 studies having on average 96 patients per trial. The placebo controlled trials of dopamine agonists, monoamine oxidase type B inhibitors, and catechol-*O*-methyltransferase inhibitors accounted for around 70% of the patients. The largest trial recruited 555 patients, but over two thirds of studies included fewer than 100 patients (median 44).

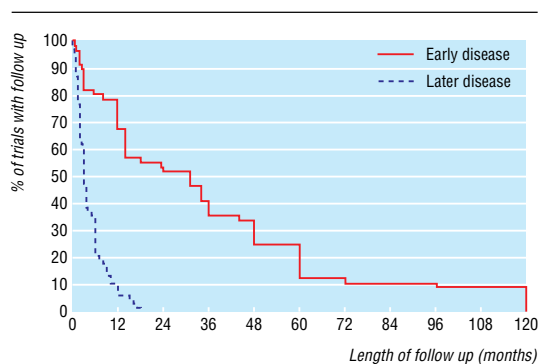
Overall, most trials were too small to produce reliable results. The largest trial accrued only 800 patients,⁵ and 60% accrued fewer than 100 patients and so had poor statistical power to detect, or refute, relatively

moderate—but nevertheless clinically worthwhile—differences between treatments. Small trials often yield false negative results (there is a real difference between treatments, but the trial is too small to show it), so beneficial treatments may be overlooked. Small trials are also more likely to produce false positive results (there is a moderate difference, or none, between arms, but by chance there seems to be a large one) as statistical significance can only be reached if the difference between treatments is implausibly big. Moreover, publication bias (trials with striking results are more likely to be published than negative ones^{6,7}) can further augment false impressions of efficacy and toxicity obtained from small trials.

Meta-analysis is a way of reducing false negative findings by increasing statistical power. It helps to reduce false positive findings by giving a more balanced view of the total evidence. An example of a false positive finding exposed by meta-analysis may be the apparent increase in mortality in patients taking selegiline in one of the largest trials of early Parkinson's disease.⁸ This reached borderline significance ($P=0.05$) but was not confirmed by a meta-analysis that includes other similar trials.⁹ Hence it seems most likely that this unexpected finding—which led to the widespread abandonment of an inexpensive and effective drug—was due simply to the play of chance. There must be many other false positive results. Over 100 published trials have evaluated drug treatments for Parkinson's disease, with many different outcome measures in each trial; a few dozen would be expected to produce findings significant at $P < 0.05$ by chance.

Duration of follow up

The average length of follow up per patient in studies of early disease was 3.8 years. However, the median length of follow up per trial was just 2.0 years, as over a third of the person years of follow up came from just two studies (UKPDRG¹⁰ and DATATOP⁵), which followed up patients for up to 10 years after randomisation (figure). Around 40% of trials in early Parkinson's disease (30% of patients) did not follow up



Proportion of trials of drug treatments for Parkinson's disease with follow up by month according to planned maximum as specified for the study

patients beyond 12 months. Studies of treatment in later Parkinson's disease had even shorter follow up. The average length of follow up per patient was only five months, and median follow up per trial was just three months. No trial had more than 18 months of follow up, less than 4% of patients had more than one year of follow up, and only a quarter of patients contributed any data beyond six months.

Most patients with Parkinson's disease survive for 15 years or longer. It is essential therefore that the impact of treatment be evaluated over the longer term. We found considerable variation in the length of follow up across trials, ranging from days to years. Moreover, not all patients were followed for the time specified in the publication, with patients either defaulting from follow up or being excluded because of failure to take their drugs or other protocol violations. Trials lasting less than five years cannot properly evaluate potential neuroprotective effects of treatments and their ability to delay the onset of motor complications. Follow up of at least five years, and ideally longer, is needed to assess reliably the long term effects of drug treatment.

Outcome measures

Most trials relied on clinician based rating scales to measure motor impairments and disability (table 2), and most used the unified Parkinson's disease rating scale (UPDRS). Many trials in early disease studied time to onset of motor complications, and trials of later disease examined improvements in "on-off time" or

Table 2 Outcome measures used in trials of drug treatment for Parkinson's disease

Outcome measure	% of trials reporting outcome measure	
	Early disease	Later disease
Clinician based rating scales*	98	83
Levodopa dose reduction or delay to use	29	54
Improvements in "on-off" time or "wearing off"	7	52
Patients' diaries	7	36
Time to development of motor complications	27	0
Quality of life measures	5	7
Economic evaluation	2	0

*These included UPDRS (unified Parkinson's disease rating scale – all stages), Hoehn and Yahr, Schwab and England, Parkinson dyskinesia scale, clinician's global impression, Columbia scale, Webster scale and NWUDS (Northwestern University disability scale).

"wearing off." Only 12 trials (involving 11% of the total number of patients) reported patients' ratings of quality of life, such as SF36 and PDQ39, and just two trials (involving 2% of the total number of patients) had an economic evaluation in the main report of the trial. None of the trials assessed the impact of treatment on the carers of patients with Parkinson's disease.

Motor impairment rating scales, used as primary outcome measures in most trials, fail to assess the impact of the disease on the whole patient. For example, a recent agonist trial found a delay in time to onset of motor complications with dopamine agonists,¹¹ but at the expense of poorer control of the symptoms of Parkinson's disease, and an increase in hallucinations—which may be more important for patients and carers than motor complications. In the absence of patient-rated quality of life assessment, the balance between these competing benefits and risks is unclear. Depression, dementia, and sleep disturbance are other common problems in Parkinson's disease, especially in its later stages.¹²⁻¹⁴ Trials should include patient rated quality of life measures, such as PDQ-39, which assess all aspects of the patient's life and are sensitive to changes considered of importance to patients but not identified by clinical ratings.¹⁵

In this era of limited resources and finite health-care budgets, it is important to assess not just clinical effectiveness but also cost effectiveness. A recent Cochrane systematic review of trials comparing "modern" dopamine agonists with bromocriptine in later Parkinson's disease found that some of the newer agonists reduce off time by 30 minutes per day.¹⁶ No other differences between the agonists were found. However, this assessment of the effects of treatment on functional ability fails to give a full insight into the relevance of the treatment on patients' global quality of life (is an extra half hour of on-time of meaningful value to the patient?) and therefore whether the additional costs of these agonists are justified. Future trials should include cost effectiveness analyses along with assessment of quality of life of the patient and their carers.

Efficacy

We also undertook a basic qualitative assessment of the results of the trials (table 3). As outcome data on efficacy were inconsistently reported, we used a qualitative scoring system to synthesise the results. This simple summary should not be seen as providing definitive evidence (more detailed quantitative syntheses of the data are available for some comparisons as Cochrane reviews^{16 17}). Reasonably good evidence of the efficacy and safety of each of the main classes of drugs is available from placebo controlled trials (though often with selective eligibility criteria—for example, only younger patients included—that limit the generalisability of the results). However, there is very little evidence on the comparative efficacy of classes of drugs. A recent review of treatments of Parkinson's disease concluded: "There are nearly no data for comparisons between interventions."¹⁸

Quality of trials

In common with most areas of medicine, trial design has improved steadily with time. However, many

Table 3 Available efficacy data in trials of drug treatments for Parkinson's disease

Comparison*	Clinical disability or impairment scales	Time to development of motor complications	Improved "on-off" time or "wearing off"	Side effects or toxicity	Reduction or delay in dose of levodopa
Early disease:					
Dopamine agonist v placebo	+++			(-)	
Monoamine oxidase type B inhibitor v placebo	+++			0	
Dopamine agonist v levodopa		+++		(-)	+
Later disease:					
Dopamine agonist v placebo	++		++	(-)	+
Monoamine oxidase type B inhibitor v placebo	++		++	(-)	+
Catechol-O-methyltransferase inhibitor v placebo	++		++	(-)	+
Catechol-O-methyltransferase inhibitor v dopamine agonist	0		0	0	0

Scores indicate comparative benefit or harm of drug shown on left: (+) = possible benefit; + = some benefit; ++ = moderate benefit; +++ = large benefit; 0 = no difference; (-) = possible adverse effect; - = adverse effect.

*Only comparisons with three or more trials and only trials comparing one class against another or placebo have been included.

deficiencies remain, such as inadequate numbers of patients, limited length of follow up, and a preoccupation with motor impairment measures of little relevance to patients and healthcare purchasers compared with quality of life and health economic outcomes. Substantial uncertainties about fundamental aspects of treating Parkinson's disease remain, and after decades of research into both early and later Parkinson's disease we still have little evidence on which to base decisions between different classes of drug.

Conclusions

It is important to determine more reliably the comparative efficacy of the classes of drug used in Parkinson's disease. Realistically, the differences between two active agents will be moderate, but nevertheless potentially important, and larger trials involving a few thousand patients are needed to detect such differences. Substantial experience indicates that large scale recruitment is best achieved with simple and pragmatic trial designs that fit in as much as possible with routine clinical practice and impose minimal extra workload on clinical staff.¹⁹ These often use factorial designs (rarely used in Parkinson's disease trials) which permit more than one question to be answered for little additional cost. Large pragmatic factorial trials in acute diseases have had a major impact on improving the treatments available for cancer (for example, QUASAR, with more than 7000 participants²⁰), heart disease (ISIS-4, n=58 052²¹), and stroke (IST, n=19 435²²). Although trials in neurodegenerative diseases do not need to be this large, those who treat patients with Parkinson's disease must accept the need for large pragmatic trials and participate in them.

Funding: This review arose from background research for the PDMED Trial, which is funded by a grant from the Health Technology Assessment Programme of the NHS to the University of Birmingham Clinical Trials Unit. The views and opinions expressed herein do not necessarily reflect those of the Health Technology Assessment Programme.

Competing interests: CEC has received fees from the manufacturers of several of the drugs discussed in this review for attending conferences, presenting lectures, and consultancy.

- Quinn NP. Parkinson's disease: clinical features. *Baillieres Clin Neurol* 1997;6:1-13.
- Clarke CE. Medical management of Parkinson's disease. *J Neurol Neurosurg Psychiatry*. 2002;72(suppl 1):22-7i.
- Oertel WH, Quinn NP. Parkinson's disease: drug therapy. *Baillieres Clin Neurol* 1997;6:89-108.
- Moher D, Schulz KF, Altman DG, for the CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 2001;357:1191-4.
- Parkinson Study Group. Mortality in DATATOP: A multicenter trial in early Parkinson's disease. *Ann Neurol*. 1998;43:318-25.
- Dickersin K, Min YI, Meinert CL. Factors influencing publication of research results. Follow-up of applications submitted to two institutional review boards. *JAMA* 1992;267:374-8.
- Dickersin K. How important is publication bias? A synthesis of available data. *AIDS Educ Prev* 1997;9(suppl 1):15-21.
- Lees AJ on behalf of the Parkinson's Disease Research Group of the United Kingdom. Comparison of therapeutic effects and mortality data of levodopa and levodopa combined with selegiline in patients with early, mild Parkinson's disease. *BMJ* 1995; 311:1602-7.
- Counsell C. Effect of adding selegiline to levodopa in early, mild Parkinson's disease. *BMJ* 1998;317:1586.
- Lees AJ, Head J, Ben-Shlomo Y on behalf of the Parkinson's Disease Research Group of the United Kingdom. Ten year follow up of three different initial treatments in de-novo Parkinson's disease. A randomised trial. *Neurology* 2001;57: 687-94.
- Rascol O, Brooks DJ, Korczyn AD, De Deyn PP, Clarke CE, Lang AE. A five-year study of the incidence of dyskinesia in patients with early Parkinson's disease who were treated with ropinirole or levodopa. *N Engl J Med* 2000;342:1484-91.
- Cummings JL. Depression and Parkinson's disease: a review. *Am J Psychiatry* 1992;149:443-54.
- Arsland D, Tandberg E, Larsen JP, Cummings JL. Frequency of dementia in Parkinson's disease. *Arch Neurol* 1996;53:538-42.
- Lees AJ, Blackburn NA, Campell VL. The nightmare problems of Parkinson's disease. *Clin Neuropharmacol* 1998;11:512-9.
- Jenkinson C, Fitzpatrick R, Peto V. The Parkinson's disease questionnaire: user manual for the PDQ-39, PDQ-8 and PDQ summary index. Oxford: Health Services Research Unit, University of Oxford, 1998.
- Clarke CE, Deane KNO, Davies P, Speller JM. Systematic review of adjuvant dopamine agonist therapy for levodopa induced motor complications in Parkinson's disease. *Parkinsonism Related Disord* 2001;7:S53.
- Clarke CE, Speller JM, Clarke JA. Pramipexole for levodopa-induced complications in Parkinson's disease. *Cochrane Database Syst Rev* 2000;(4):CD002261.
- Rascol O, Goetz C, Koller W, Poewe W, Sampaio C. Treatment interventions for Parkinson's disease: an evidence based assessment. *Lancet* 2002;359:1589-98.
- Collins R, Peto R, Gray R, Parish S. Large-scale randomised evidence: trials and overviews. In: Weatherall D, Ledingham JGG, Warrell DA, eds. *Oxford Textbook of Medicine*, 3rd ed. Oxford: Oxford University Press, 1996:21-32.
- QUASAR Collaborative Group. Comparison of fluorouracil with additional levamisole, higher-dose folic acid, or both, as adjuvant chemotherapy, for colorectal cancer. *Lancet* 2000;335:1588-96.
- ISIS-4 (Fourth International Study of Infarct Survival) Collaborative Group. ISIS-4: a randomised factorial trial assessing early oral captopril, oral mononitrate, and intravenous magnesium sulphate in 58,050 patients with suspected acute myocardial infarction. *Lancet* 1995;345:669-85.
- International Stroke Trial Collaborative Group. The international stroke trial (IST): a randomised trial of aspirin, subcutaneous heparin, both, or neither among 19,435 patients with acute ischaemic stroke. *Lancet* 1997;349:1569-81.