Paper presented to the Health Economists Study Group Meeting,
University College Cork, 23-25 June 2010

A systematic review of decision-analytic models and critical appraisal of modelling methodology in Parkinson's disease

James Shearer, Colin Green

Collaborators: Dave Wright, Paul Thompson, John Zajicek, Jeremy Hobart, Carl Counsell, Craig Ritchie, Gordon Wilcock

Peninsula College of Medicine and Dentistry, University of Exeter
(james.shearer@pms.ac.uk)

Abstract

Aims: To review the methods available to model the cost-effectiveness of interventions in Parkinson's disease (PD), and to consider future directions that may improve methods for studies using modelling frameworks to model progression of PD over time.

Methods: A systematic search of the health care literature up to October 2009 to identify model-based evaluations in PD. The methods, structure and data used in the identified modelling studies were appraised using good practice guidelines in decision-analytic modelling.

Results: The review identified 17 evaluations using modelling methods in PD. Most studies evaluated medications (n=12), three evaluated screening technologies, and two examined surgical procedures. There were no models for other types of interventions such as rehabilitation or behavioural therapy. Most models relied on motor symptoms to characterise disease progression and severity. No models evaluated interventions targeted towards non-motor symptoms such as neuropsychiatric problems (cognition, mood, behaviour), or autonomic dysfunction (gastrointestinal problems, genitourinary problems).

Conclusions: The decision-analytic models took a narrow view of PD and its impact on the lives of patients, their health care needs and costs. Reliance on motor symptoms to model long term effects of PD treatment was a limitation in models as it is now well-accepted that non-motor symptoms are a major part of PD disease progression and have large potential economic impacts on costs and outcomes. More comprehensive models of disease progression including both motor and non-motor symptoms are needed.

Introduction

Parkinson's disease (PD) is a chronic neurodegenerative disorder affecting around 1% of the population aged >60 years worldwide [1]. The progressive course of PD is associated with increasing morbidity and mortality. Disease progression in PD is caused by ongoing degeneration of dopamine neurons and is characterised by symptoms of resting tremor, rigidity, bradykinesia (slowness of movement) and postural instability [2]. The natural history of PD is also characterised by an increasing incidence of non motor symptoms including dementia, depression, psychosis, sleep disorders and autonomic dysfunction [3]. There are currently no interventions that stop the lifelong progression of PD. PD motor symptoms are initially well controlled by medications which enhance dopaminergic function, however, long-term use of these medications is associated with motor fluctuations and dyskinesias that eventually add to the burden of PD. Non-pharmacological approaches are also used to manage the motor and non motor symptoms of PD including surgical stimulation of brain neurones, rehabilitation, physiotherapy, occupational therapy, speech therapy, and behavioural and psychological interventions [4].

As populations age, PD is expected to represent an increasingly significant health care burden on individuals and health care systems. The physical disability, morbidity and psychosocial difficulties associated with disease progression in PD negatively impact the health-related quality-of-life (HRQL) of people with PD. Against the growing demand for health care and support from those with PD, it is inevitable that future health technology developments (as well as existing health technologies) for PD will be subject to scrutiny as health care resources become increasingly constrained and subject to difficult priority-setting conditions. Health care decision makers will be faced with difficult decisions over the funding (or not) of health care for PD, and it will be important to establish the cost-effectiveness of health care interventions which aim to improve diagnosis, prognosis, symptom control, or modify disease progression.

Cost-effectiveness analysis is now common place in decision-making (e.g. health technology appraisal), however, the longer-term empirical data needed to inform health care decisions for PD (such as home care requirements, institutionalisation, patient quality of life, models of disease progression, morbidity and mortality) are rarely provided by clinical trials or observational studies. In the absence of empirical data, modelling methods have been used to estimate the future costs and future consequences of alternative treatment strategies. Where data are not available on longer term outcomes decision-analytic models provide a logical mathematical framework to extrapolate intermediate outcomes (clinical endpoints, resource use) from a range of sources including clinical trials, to final outcomes relevant to decision makers [5][6].

The literature on decision-analytic modelling in PD is relatively undeveloped. Two recent reviews, one of decision-analytic models in PD [7] and another of the cost effectiveness of pharmacotherapies in early PD [8] found only eight published studies. Both found considerable methodological differences between the published models which

impeded any comparison of findings. In addition Siebert et al. [7] found that models did not adequately consider the full spectrum of PD symptoms particularly those that considered drug-related motor complications. The aim of the current paper is to present a summary review of the literature on model-based evaluations in PD, in order to evaluate the strengths and limitations of decision-analytic models in PD, and to assess the need for new or improved modelling methods for the economic evaluation of interventions in PD.

Methods

A systematic search of the health care literature was undertaken to identify models used in evaluations for PD, published in English up to October 2009. MEDLINE and EMBASE databases were searched using a search strategy designed to identify decision-analytic models that had been used to model costs and/or consequences of interventions for PD over time. An example of the search terms used for the MEDLINE search appears in Appendix 1. The search terms were adapted as necessary to include searches in EMBASE, the HTA Database, NHS EED, CEA Registry, PsycInfo and EconLit databases. The search produced a total of 561 citations which were reviewed separately by CG, JS to identify relevant papers. Complete references were obtained for 61 citations which were identified as potentially relevant and these were further reviewed to identify studies for inclusion. Reviews, cost studies, methodological papers, outcome studies and epidemiological models which were not evaluations were excluded. Economic evaluations conducted alongside clinical trials were excluded unless they also modelled PD over time. Reference lists and citations of retrieved articles and existing reviews were further checked to ensure that no eligible studies had been missed. The included modelling studies were reviewed against guidelines for good practice in decision-analytic modelling [9]. The results of the model appraisals is summarised in narrative and tabular form principally under the broad criteria of structure, data and uncertainty/consistency.

Results

Description of included models

The literature review identified 17 studies using decision-analytic models to estimate the impact of interventions for PD on long-term costs and outcomes. The aim of nearly all the models was to evaluate the cost-effectiveness of PD interventions, with one example of a cost-minimisation analysis [10]. Details of the models including objectives, perspectives and findings are summarised in Table 1. All of the included models were concerned with treatments that aimed to manage motor symptoms (resting tremor, rigidity and bradykinesia), postural instability or the minimisation of treatment-induced motor complications (motor fluctuations, dyskinesias). Twelve models evaluated medications used either in place of levodopa[1] in early PD, or combined with levodopa in later stages of the disease [10-21]. Comparisons were based on how well medications

[1] Levodopa, which is converted into dopamine in the brain, is the first line treatment to control the motor symptoms of PD. Unfortunately, prolonged use of levodopa is associated with progressively worsening side effects.

controlled PD motor symptoms measured using clinical rating scales and also the relative incidence of medication-related motor complications (i.e. the proportions of patients with/without complications or proportions of the day patients experience motor fluctuations referred to as 'off time'. Three models [22-24] evaluated screening technologies, and two modelled the impact of surgical interventions in late stage patients or those who could not the side-effects of medication [25, 26]. None of the identified models examined interventions for rehabilitation, psychosocial therapies or other interventions targeted towards non-motor symptoms.

Fourteen of the included models were Markov state transition models with health states defined by clinical rating scales of motor functioning or the incidence/prevalence of medication-related motor fluctuations. Disease progression was modelled by moving cohorts of patients between health states defined by PD severity, treatment setting and/or the presence of medication-related motor complications. Seven models [13-15, 17, 19, 20, 25] based health states on Hoehn Yahr (HY) stages, a commonly used clinical system for rating PD severity, which has five stages characterised by motor symptoms, postural instability and loss of mobility. Two studies [21, 26] modelled disease progression using the UPDRS, another commonly used, comprehensive measure of motor functioning. Five models [12, 16, 18] defined their health states according to the absence/presence of treatment-related motor complications. Two diagnostic screening models [22, 23] defined health states according to the accuracy of diagnosis and subsequent adequate/inadequate treatment. The remaining three models consisted of decision tree analyses [10, 24] and a mathematical model based on a simple linear conceptualisation of disease progression using the UPDRS [21]. Fourteen models factored in the impact of mortality, either specific PD mortality [12, 13, 15, 20-22, 25, 26] or the age/gender matched population [14, 17-19, 23, 27], with all assuming no difference between treatments. The time horizons used in the models varied from one year to a lifetime perspective with most (12/17) based on 5 to 10 years. Cycle length varied between four and 12 months. Only one out of the six models using a 12 month cycle reported the use of half-cycle corrections.

The choice of analytical perspective usually determined which costs were included in analyses although there were a number of important differences in the cost components even in those models that took a similar analytical perspective. Five models stated a societal analytic perspective [17-19, 21, 26], seven took a third party payer (TPP) perspective either UK NHS or other national health payer [11, 14, 15, 20, 22-24] and four models presented results from both perspectives [10, 12, 13, 16]. Models that reported results from a TPP perspective generally only considered direct health and social care costs. Models that presented analyses from a societal perspective also included informal care costs, nursing home costs, costs from lost productivity, and indirect health care costs. The costs of informal care by friends and family were included in five analyses from the societal perspective [10, 12, 15, 16, 19] although informal care costs were explicitly excluded in two other models [18, 21] with Nuijten et al citing a lack of reliable data. Four models included costs from lost productivity and early retirement [10, 12, 15, 21] with most of these using the human capital method to calculate costs of lost income and productivity. Other models specifically excluded lost productivity to avoid potential

double counting from utility valuations [19, 26], while others cited the fact that the average age of patients was above retirement age [16, 18]. Long term residential or nursing home care was included in four models [10, 20, 21, 26]. The costs of informal care, institutional care and lost productivity usually had a significant impact on cost-effectiveness when included in models.

Ten models [11-13, 16-19, 21, 25, 26] reported QALYs as their principal health outcome. Health state utility values were obtained from a number of different sources in different settings based on different valuation techniques. This yielded a wide range of values attributed to HY stage, changes in UPDRS scores or the presence/absence of medication-related side effects. Nine of the ten models derived their utility values from four published studies [28-31], with Shimbo et al [19] generating their own values from a concurrent survey of 470 Japanese PD patients. The other sources of utility data used in the models were based on smaller samples (between 60 and 258 patients) of non-institutionalised patients mainly in the early stages (HY 1-3) of PD limiting their value to models concerned with later or more severe stages of the disease. Health state values for later disease stages (i.e. HY 4/5) were sometimes based on small samples of 3 or 4 patients [31]. Four models [11, 12, 16, 18] concerned with the specific impact of medication-related motor fluctuations or periods of the day spent in 'off-time', obtained utility values from a survey of patient preferences on control of PD symptoms and motor fluctuations ('off-time') in 60 US PD levodopa patients [27]. Seven models did not use QALYs as their main health outcome citing a lack of data about the impact of HRQL on specific aspects of PD and its treatment such as the consequences of misdiagnosis [22-24] or medication-related side effects [10, 14, 15].

Appraisal of the included models

The critical appraisal was guided by a framework for quality assessment and best practice in decision-analytic modelling developed by Philips et al [9]. This framework assesses dimensions of quality and good practice under the general themes of model structure, data and consistency. Philips et al present 15 quality criteria, each with suggested questions for critical appraisal of models, generally checking whether each criterion was clearly stated, consistent with related criteria, justified with evidence and/or considered reasonable or appropriate. The models included in the present review were assessed against each of the Philips et al criteria and selected findings are summarised in Table 2. A simplified presentation is used here to draw attention to areas of strength and weakness in the published models. Only those criteria where there were questions or concerns about the performance of models are included in Table 2, in order to focus on potential areas for improvement in modelling methodology, although areas of strength in the models are acknowledged in the text.

Structure

Choices over model structure represent the first step in the development of a decision-analytic model. The structure of a model determines how the various inputs (costs, effectiveness, health states, utilities) are brought together to predict the final outcomes of interest to decision-makers (QALYs, long term outcomes, health events, marginal health gains and costs). The guidelines developed by Philips et al suggest that models should 1) describe their structure clearly, 2) provide a justification or evidence for the structure, and 3) ensure that the structure reflects disease progression in PD and the impact of treatment. Nine components of model structure are specifically appraised under the Philips et al framework including: the statement of the decision problem; the statement of the model scope/perspective; the rationale for the model structure; structural assumptions; strategies/comparators; model type; time horizon; disease states/pathways; and cycle length.

A number of structural issues were clearly presented in almost all the models including statements of the decision problem, model objectives and scope, analytical perspective, the choice of comparators, and the choice of model type. Most models also clearly described components relating to time including time horizon, cycle length, treatment duration and duration of treatment effect. Most justified their choice of time horizon and cycle length based upon the availability of clinical trial data or cohort data, model objectives (i.e. time needed to confirm an uncertain diagnosis), the age of patients or their prognosis (i.e. time to end-stage disease). Some models extrapolated effects from short term clinical trials (6-12 months) to longer time horizons (5-10 years) assuming a constant treatment effect over time [11-13, 17, 20]. Other models, however, took a more conservative approach and only used the duration of treatment effect established in clinical trials after which disease progression in treated patients reverted to baseline [16, 18, 19, 21, 25, 26]. Three models were able to rely on treatment effects observed in long term (5 year) controlled clinical trials [10, 14, 15].

The rationale for the structure used in models should be based on a coherent theory of disease progression in PD and its treatment [5, 9, 32]. Nine models [13-15, 17, 19-21, 25, 26] used progression in clinical measures of motor functioning as a proxy for disease progression in PD. The models justified this on the basis that measures such as the UPDRS and HY stages were well accepted, reliable and validated instruments widely used in clinical practice and clinical trials. The most common assumption was that disease progression and any subsequent impact from treatment was characterised by linear [11, 15, 19, 21, 25] and in one case exponential [26] declines in clinical measures of motor function although no evidence was given to support this assumption.

Seven models [13-15, 17, 19, 20, 25] relied on health states based on the HY clinical classification system for PD severity stage. The main advantage of HY stages are that they reflect relevant clinical milestones in PD starting with unilateral (one-sided) motor impairment (stage 1), to bilateral (both sides affected) (stage 2), postural impairment (stage 3), to the final stages of loss of independence (stage 4) and loss of mobility (stage5). Two models specifically cited evidence that patient costs and utilities were

related to disease progression according to HY stage [13, 17]. A key limitation of HY stages, however, is that it exclusively focuses on motor and postural reflex impairment and does not reflect the full spectrum of PD including non motor symptoms [33]. Measuring progression of motor symptoms alone does not sufficiently describe PD progression [34] and changes in treatment-related motor complications may not reflect changes in underlying biological progression of PD [7].

Two models [21, 26] used changes in UPDRS scores to measure disease progression. The UPDRS is a comprehensive measure of overall motor function, however, there is no agreement about how UPDRS scores correspond to disease stages (i.e. early, mild, moderate, severe, late) [35], or what score constitutes a minimum clinically significant effect size [35]. As with HY stages, the focus of UPDRS is on motor impairment rather than non motor problems. There was also no evidence given to support the assumption that changes in UPDRS are related to changes in other outcomes such as costs or quality of life.

Seven models wholly or partially based their structure on the incidence of medication-related motor complications rather than progression of the underlying PD [10-12, 14-16, 18]. Three models incorporated motor complications either into health states defined by either HY stage [15] or levodopa treatment [10, 11] with or without motor complications, or as a distinct absorbing health state [14]. Outcomes were based on either the proportion of complication-free patients [10, 15] or years of avoided motor complications [11, 14]. Three other models used a similar structure with health states defined using a threshold of 25% for the proportion of day affected by levodopa-related motor fluctuations. The threshold was based on a German cost study [28] which found a significant decrease in utility and higher costs associated with having motor fluctuations for more than 25% of a day. One of these models [12] acknowledged that factors affecting disease progression in PD had not been considered, and accordingly the time horizon of their model was restricted to 2 years to reflect the low 'external validity' of projections beyond that point.

Data

The methods used to identify, transform and incorporate data into models are an important aspect of modelling methodology although the process can be mathematically complex. This section uses the Philips et al quality framework to assess whether the included models identified, measured and valued relevant cost and outcome data in an appropriate and transparent way. Philips et al consider data issues under three broad sections: data identification; pre-model analysis and data incorporation. Data identification refers to the methods used to identify data inputs and includes an assessment of the quality and the reliability of the data. Pre-model analysis refers to the statistical methods used to transform data into a form that can be used in models, for example from rates into transit probabilities, or from intermediate clinical rating scales such as the UPDRS into longer term health outcomes. Philips et al framework considers pre-model analysis issues under three subsections: baseline data (often relating either to natural disease progression or an experimental control group); the effects of treatment and quality-of-life weights. Data incorporation refers to the description of the data and its

sources, and in what manner the data has been incorporated into the model (i.e. as point estimates or distributions).

The identification and incorporation of data in most models was assessed as satisfactory, i.e. most studies provided references, a description of the input data and sources and a justification of the appropriateness and validity of the data inputs. Very few models described systematic literature reviews [18, 20, 21, 24, 26] although this was mitigated by the relatively small size of the literature around costs and outcomes particularly for new interventions. The majority of models were able to use local cost [11, 13-15, 17, 21, 24, 25] and utility [13, 15, 17, 19, 21] data specific to their decision-problem thereby enhancing the validity of their analyses. However, some models needed to rely on expert opinion to provide data estimates [16, 19, 20, 22, 23, 26] and only few provided a description of the process used to elicit the views of experts, such as Delphi Panels [16, 22, 23]. A lack of direct empirical evidence around resource use and costs associated with dyskinesias and motor fluctuations ('off-time') led some models [11, 12, 14] to apply questionable ratios to distinguish costs between health states with and without motor complications which were not varied in sensitivity analysis. In other models the way in which data from multiple clinical sources was combined to produce point estimates for effectiveness was not always clear [21].

Pre-model data analysis

The way models combine or convert different types of data drawn from different sources into data inputs that can be incorporated into models is important. Some of the identified models were not clear about the statistical methods or other calculations used in pre-modelled data. The calculation of transit probabilities and extrapolation assumptions are the central engines of a Markov model and need to be clearly stated and justified. The methods used to derive annualised transit probabilities from cited data sources were not clear or consistent in some models [11, 15] (Smala et al., Haycox et al.,). Smala et al [15] calculated annualised transit probabilities using data from a cohort of 330 patients (male 77%, mean age PD onset 58) followed between1968 to 1994 whose progress after starting levodopa was followed through each HY stage until they died [36]. The method used to generate annualised transit probabilities from the mean survival intervals reported in months from the cited data source, however, was not explained and original data were not provided. Haycox et al [11] appeared to have used a constant annual transit probability calculated from a controlled trial to model movement from pramipexole (the comparator medication) to levodopa over a five year time horizon, but observed non-linear transit probabilities from an uncontrolled study extension for the experimental rasagiline group. The approach was justified on the basis that only four years of uncontrolled observational data were available for the comparator group.

Models took different approaches to how treatment impacted disease progression. The impact of treatment on the course of PD was assumption and not empirically based in some models [19, 21, 25, 26] often based on a constant linear effect over time after a one-off improvement in HY stage or drop in UPDRS score. Other models [11-13, 17, 20] extrapolated a constant treatment effect beyond the period of the clinical trial, only one

again these assumptions were rarely tested in sensitivity analysis. Those that were concerned about differences in the side-effects of different treatment assumed there no effect of treatment on disease progression despite the evidence that levodopa exerts superior control over PD motor symptoms compared to alternative medications [37].

Generally utility data sources (health state values) were referenced and appropriate. Four models [11, 12, 16, 18] that used the same study described previously [27] for utility values for patients with and without motor complications selected values depending on whether they were produced using a VAS technique or standard gamble. The VAS results were more widely dispersed across the health states (i.e. the differences between health state values were not so great using standard gamble). Palmer et al [16] and Nuitjen et al [18] used the higher values generated using the standard gamble techniques arguably representing the most conservative approach. Haycox et al [11] used the lower VAS scale values which they varied by +10% and -20% in sensitivity analysis rather than using the larger differences between standard gamble and VAS methods. Hudry et al [12] reported a weighted average of the values from the Palmer et al study [27] but how these were calculated was not presented.

Uncertainty and consistency

Uncertainty is inherent in decision analysis and it is recommended that decision-analytic models should conduct sensitivity analysis around methodological, structural, heterogeneity and parameter uncertainty [9]. While all the included models reported sensitivity analysis most involved very limited one-way deterministic tests of selected sources of parameter and methodological/structural uncertainty. Uncertainty was handled in a selective manner in some models and authors and reviewers should ensure that where inputs are largely assumption-based or based on expert opinion these are subject to rigorous sensitivity analysis. Only two models [13, 17] presented probabilistic sensitivity analysis, which has been recommended because it handles the joint effect of uncertainty over key parameters. Methodological and structural uncertainty was generally dealt with in an outline manner (i.e. by varying discount rates, time horizon). Changes in parameter valuation methods or the way disease progression was characterised were rarely examined. The impact of uncertainty associated with patient heterogeneity was not examined in any of the models although this might reflect the lack of reliable data about identifiable, systematic sub-group differences. None of the models reported evidence for internal consistency (the underlying mathematical logic) or validity, apart from claims of expert review [18, 24]. There was some consideration of external validity by comparing results with other models but this was necessarily limited given the small size of the relevant literature.

Discussion

The 17 decision-analytic models identified in this systematic review used a number of often novel and sophisticated approaches to provide the long term projections of costs and outcomes needed for health care decision-making in PD using very limited sources of clinical, epidemiological and economic data. We did not set out to compare the cost-effectiveness of interventions but were more interested in documenting developments in the models particularly on issues related to modelling disease progression across the full spectrum of PD. Our systematic appraisal of the reporting of methods used in the models found that they performed well in terms of stating issues such as the decision problem, analytical perspective, justifying time horizon, the discussion of the identification and incorporation of input data, and exploring uncertainty relating to model parameters. Models were also transparent about the rationale and assumptions underlying their chosen structures. However, models performed less well when it came to providing the evidence to support decisions about model structure or the synthesis of data inputs. None of the models considered alternative model structures, and only a minority attempted to explicitly justify the selection of model structure or consider alternative outcomes or theories of disease progression. None of the models provided evidence of internal or external model validation. The most important limitation in the model-based evaluations identified in this review was the exclusive focus on motor functioning to model the impact of treatment on disease progression in PD.

The natural history of PD is characterised by an increasing prevalence of a large number of non-motor symptoms that have significant impact on the HRQL of patients and the costs of their health care [34]. These include depression, dementia, sleep disorders, behavioural problems, gastrointestinal problems, and genitourinary problems. Depression is associated with advancing disease severity and significantly impacts the HRQL in a large proportion of patients [38, 39] while behavioural problems and dementia significantly predict the need for nursing home placement [40]. None of the included models evaluated interventions targeted towards non-motor symptoms in PD even though development of non motor symptoms may provide a better reflection of disease progression in PD than motor symptoms [41]. Van Rooden et al [41] conducted a factor analysis based on the full spectrum of motor and non-motor symptoms in a cohort of 397 (HY median stage 2) in order to identify which factors best reflected disease severity and progression. The strongest factor comprised most of the non-motor domains including cognitive impairment, autonomic dysfunction, psychosis, and postural instability with smaller contributions from daytime sleepiness and depression. The weakest factor comprised the classical PD symptoms of tremor, bradykinesia and rigidity although weaker correlations in this factor may have been due to a masking effect of medication (70% were receiving dopamine agonists and subjects were assessed while they were in the 'on' state). The authors suggested that elements in the first factor could provide a new basis for monitoring disease progression in PD. These preliminary findings emphasised the limitations in measures based on motor sign (HY Staging, UPDRS) to predict disease progression.

Although there are currently no treatments proven to slow or reverse disease progression in PD this is an area of intense research effort [42]. None of the modelling approaches to date would be capable of making the distinction between the disease modifying and the symptomatic effects of interventions. Only the model proposed by Hjemlgren et al [25] specifically aimed to model costs and benefits of disease-modifying interventions and this was limited because it relied solely on motor functioning as assessed by HY stage. New methods/ways of measuring disease severity and progression (change in disease severity over time) are needed if decision-analytic models are to be able to extend to evaluate interventions which aim to modify disease progression. Future models could benefit from insights gained from pharmacological modelling studies which have tackled the issue of disentangling disease-modifying from the symptomatic effects of new pharmacotherapies [43]. Treatments may have symptomatic effects that do not change the rate of disease progression and protective effects which slow disease progression or a combination of both types of effects. How these effects are incorporated into a disease progression model is important. Economic models which do not recognise this may overestimate treatment benefits and give incorrect estimates of cost-effectiveness. Various statistical and experimental techniques have been proposed to disentangle these effects although all are limited to motor function, usually measured by the UPDRS [43-46].

Conclusion

Reliance on motor symptoms to model long term effects of PD treatment was a limitation in models as it is now well-accepted that non-motor symptoms are a major part of PD disease progression and have large potential economic impacts on costs and outcomes. More comprehensive models of disease progression including both motor and non-motor symptoms are needed. Better measures and analytical approaches are needed for models to reflect the impact of interventions on the complete spectrum of PD. Such instruments and approaches will need, at a minimum, to include non motor symptoms, be patient-centred and reflect best-practice in terms of their technical development. Events and health states in future models need to be defined in terms of both non motor and motor symptoms. The lack of models that project the long term costs and consequences of interventions other than medications or medical procedures was also a serious shortcoming in the literature.

References

1. Samii, A., J. Nutt, and B. Ransom, *Parkinson's disease.* Lancet, 2004. **363**: p. 1783-1793.
2. Ibbotson, T. and K. Goa, *Management of Parkinson's disease: defining the role of entacapone.* Disease Management and Health Outcomes, 2002. **10**(10): p. 643-659.
3. Chaudhuri, K., D. Healy, and A. Schapira, *Non-motor symptoms of Parkinson's disease: diagnosis and management.* Lancet Neurology, 2006. **5**: p. 235-245.

4. National Collaborating Centre for Chronic Conditions, *Parkinson's disease: National clinical guideline for diagnosis and management in primary and secondary care* 2006, Royal College of Physicians: London.
5. Weinstein, M., et al., *Principles of good practice for decision-analytic modeling in health-care evaluation: report of the ISPOR Task Force on good research practices - modeling studies.* Value in Health, 2003. **6**(1): p. 9-17.
6. Briggs, A., K. Claxton, and M. Sculpher, *Decision Modelling for Health Economic Evaluation*. Handbooks in Health Economic Evaluation. 2006, Oxford: Oxford University Press.
7. Siebert, U., et al., *Systematic assessment of decision models in Parkinson's disease.* Value in Health, 2004. **7**(5): p. 610-626.
8. Eggert, K., et al., *Cost effectiveness of pharmacotherapies in early Parkinson's disease.* CNS Drugs, 2008. **22**(10): p. 841-860.
9. Philips, Z., et al., *Good practice guidelines for decision-analytic modelling in health technology assessment: A review and consolidation of quality assessment.* Pharmacoeconomics, 2006. **24**(4): p. 355-371.
10. Iskedjian, M. and T. Einarson, *Cost analysis of ropinerole versus levodopa in the treatment of Parkinson's disease.* PharmacoEconomics, 2003. **21**(2): p. 115-127.
11. Haycox, A., et al., *Cost effectiveness of rasagiline and pramipexole as treatment strategies in early Parkinson's disease in the UK setting: an economic Markov model evaluation.* Drugs and Aging, 2009. **26**(9): p. 791-801.
12. Hudry, J., et al., *Cost-utility model of rasagiline in the treatment of advanced Parkinson's disease in Finland.* The Annals of Pharmacotherapy, 2006. **40**: p. 651-657.
13. Findley, L., et al., *Cost-effectiveness of levodopa/carbidopa/entacapone (Stalevo*) compared to standard care in UK Parkinson's disease patients with wearing-off.* Current Medical Research and Opinion, 2005. **21**(7): p. 1005-1014.
14. Lindgren, P., B. Jönsson, and J. DuChane, *The cost-effectiveness of early cabergoline treatment compared to levodopa in Sweden.* European Journal of Health Economics, 2003. **4**: p. 37-42.
15. Smala, A., et al., *Cabergoline versus levodopa monotherpay: a decision analysis.* Movement Disorders, 2003. **18**(8): p. 898-905.
16. Palmer, C., et al., *Cost effectiveness of treatment of Parkinson's disease with entacapone in the United States* PharmacoEconomics, 2002. **20**(9): p. 617-628.
17. Linna, M., et al., *Probabilistic sensitivity analysis for evaluating cost-utility of entacapone for Parkinson's disease.* Expert Review of Pharmacoeconomics Outcomes Research, 2002. **2**(2): p. 91-97.
18. Nuijten, M., et al., *Cost-effectiveness analysis of entacapone in Parkinson's disease: a Markov process analysis.* Pharmacoeconomics, 2001. **4**(4): p. 316-328.
19. Shimbo, T., et al., *Cost-effectiveness analysis of dopamine agonists in the treatment of Parkinson's disease in Japan.* PharmacoEconomics, 2001. **19**(8): p. 875-886.
20. Davey, P., et al., *Cost-effectiveness of pergolide compared to bromocriptine in the treatment of Parkinson's disease: a decision analytic model.* Value in Health, 2001. **4**(4): p. 308-315.

21. Hoerger, T., et al., *Cost effectiveness of pramipexole in Parkinson's disease in the US.* Pharmacoeconomics, 1998. **14**(5): p. 541-557.
22. Van Laere, K., et al., *The cost effectiveness of ^{123}I-FP-CIT SPECT imaging in patients with an uncertain clinical diagnosis of parkinsonism.* European Journal of Nuclear Medical amd Molecular Imaging, 2008. **35**: p. 1367-1376.
23. Antonini, A., et al., *Cost-effectiveness of ^{123}I-FP-CIT SPECT in the differential diagnosis of essential tremor and parkinson's disease in Italy.* Movement Disorders, 2008. **23**(15): p. 2202-2209.
24. Dodel, R., et al., *Dopamine transporter imaging and SPECT in diagnostic work-up of Parkinson's disease: a decision-analytic approach.* Movement Disorders, 2003. **18**(Suppl 7): p. S52-S62.
25. Hjelmgren, J., et al., *Estimating the value of novel interventions for Parkinson's disease: an early decision-making model with application to dopamine cell replacement.* Parkinsonism & Related Disorders, 2006. **12**: p. 443-452.
26. Tomaszewski, K. and R. Holloway, *Deep brain stimulation in the treatment of Parkinson's disease: a cost-effectiveness analysis.* Neurology, 2001. **57**: p. 663-671.
27. Palmer, C., et al., *Patient preferences and utilities for 'off-time' outcomes in the treatment of Parkinson's disease.* Quality of Life Research, 2000. **9**: p. 819-827.
28. Dodel, R., K. Berger, and W. Oertel, *Health-related quality of life and health care utilisation inpatients with Parkinson's disease: impact of motor fluctuations and dyskinesias.* Pharmacoeconomics, 2001. **19**: p. 1013-1038.
29. Chrischilles, E., et al., *The health burdens of Parkinson's disease.* Movement Disorders, 1998. **13**(3): p. 406-413.
30. Keränen, T., et al., *Economic burden and quality of life impairment increase with severity of PD.* Parkinsonism & Related Disorders, 2003. **9**(3): p. 163-168.
31. Schrag, A., M. Jahanshahi, and N. Quinn, *How does Parkinson's disease affect quality of life? A comparison with quality of life in the general population.* Movement Disorders, 2000. **15**: p. 1112-1118.
32. Sculpher, M., E. Fenwick, and K. Claxton, *Assessing quality in decision analytic cost-effectiveness models: a suggested framework and example of application.* Pharmacoeconomics, 2000. **17**(5): p. 461-477.
33. Goetz, C., et al., *Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: Status and recommendations The Movement Disorder Society Task Force on rating scales for Parkinson's disease.* Movement Disorders, 2004. **19**(9): p. 1020-1028.
34. Poewe, W., *The natural history of Parkinson's disease.* Journal of Neurology, 2006. **253**(Suppl 7): p. 2-6.
35. Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease, *The Unified Parkinson's Disease Rating Scale (UPDRS): Status and recommendations.* Movement Disorders, 2003. **18**(7): p. 738-750.
36. Di Rocco, A., et al., *Parkinson's Disease: Progression and Mortality in the L-DOPA Era*, in *Advances in Neurology*, L. Battistin, et al., Editors. 1996, Lippincott-Raven Publishers: Philadelphia. p. 3-11.
37. Le Witt, P., *Levodopa for the treatment of Parkinson's Disease.* New England Journal of Medicine, 2008. **359**(23): p. 2468-2476.

38. Dowding, C., C. Shenton, and S. Salek, *A review of the health-related quality of life and economic impact of Parkinson's disease.* Drugs and Aging, 2006. **23**(9): p. 693-721.
39. Schrag, A., M. Jahanshahi, and N.P. Quinn, *What contributes to quality of life in patients with Parkinsons's disease?* Journal of Neurology, Neurosurgery & Psychiatry, 2000. **69**: p. 308-312.
40. Goetz, C. and G. Stebbins, *Risk factors for nursing hom eplaceemnt in advanced Parkinson's disease.* Neurology, 1993. **43**: p. 2227-2229.
41. van Rooden, S., et al., *Patterns of motor and non-motor features in Parkinson's disease.* Journal of Neurology, Neurosurgery & Psychiatry, 2009. **80**: p. 846-850.
42. Schapira, A.H.V., *Neurobiology and treatment of Parkinson's disease.* Trends in Pharmacological Sciences, 2009. **30**(1): p. 41-47.
43. Holford, N., et al., *Disease progression and pharmacodynamics in Parkinson Disease - evidence for functional protection with levodopa and other treatments.* Journal of Pharmacokinetics and Pharmacodynamics, 2006. **33**(3): p. 281-311.
44. Post, T., et al., *Disease system analysis: Basic disease progression models in degenerative disease.* Pharmaceutical Research, 2005. **22**(7): p. 1038-1049.
45. Chan, P. and N. Holford, *Drug treatment effects on disease progression.* Annual Review of Pharmacology and Toxicology, 2001. **41**: p. 625-659.
46. Guimaraes, P., et al., *Non-linearity of Parkinson's disease progression: Implications for sample size calculations in clinical trials.* Clinical Trials, 2005. **2**: p. 509-518.

Appendix 1: Example of Medline search and results

Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) <1950 to Present>
Search Strategy:

1	exp Parkinson Disease/
2	exp Parkinsonian Disorders/
3	1 or 2
4	parkinson$.tw.
5	3 or 4
6	economic model$.ti.
7	economic model$.ab.
8	Markov chains/
9	markov$.ti.
10	markov$.ab.
11	monte carlo method/
12	monte carlo.ti.
13	monte carlo.ab.
14	models economic/ or model, statistical
15	exp Decision Theory/
16	(decision$ adj2 (tree$ or analy$ or model$)).ti.
17	(decision$ adj2 (tree$ or analy$ or model$)).ab.
18	(natural history adj2 model$).mp.
19	(disease adj2 model$).mp.
20	6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19
21	5 and 20
22	limit 21 to English language
23	22 not animal.mp.

Table 1: Summary of characteristics of the included models

Study authors, year, country	Objective	Target population	Statistical/ analytical Approach	Health outcomes and costs	Time horizon (cycle length)	Perspective	Findings	Competing interests*
Medications								
Haycox et al., 2009 UK	CEA of rasagiline vs pramipexole	Early PD levodopa-naïve patients	Markov cohort simulation based on transit probabilities derived from trials	Time to levodopa, Time to dyskinesia, QALYs, Direct healthcare costs	Five years (six-months)	TPP	Rasagiline dominated pramipexole	Yes
Hudry et al., 2006 Finland	CEA of rasagiline+levodopa vs entacapone+levodopa vs levodopa	PD patients with severe motor fluctuations	Markov cohort simulation based on transit probabilities derived from trials	Time with ≤25% off-time/day, QALYs, Mortality, Direct health care costs, Informal home care Early retirement	Two years (four-months)	Societal and TPP	Both combination therapies dominated levodopa alone	Yes
Findley et al., 2005 UK	CEA of entacapone+levodopa vs levodopa	PD patients with severe motor fluctuations	Markov cohort simulation based on transit probabilities derived from trials	Time in modified HY stage, Mortality, QALYs, Direct health care costs, Direct social services costs, Direct private patient costs	Ten years (six-months)	Societal and TPP	Entacapone was cost saving from societal perspective, cost effective from the TPP perspective	Yes
Iskedjian and Einarson, 2003 Canada	CMA of ropinerole vs levodopa	Early stage PD patients without dyskinesias	Decision tree cohort simulation with event probabilities derived from a trial and expert opinion	Proportion of patients with dyskinesias, Direct health care costs, nursing home admissions, informal caregiver time,	Five years (Not stated)	Societal and TPP	Ropinerole was cost saving from a societal perspective, and associated with a small incremental TPP cost.	Yes

				Lost productivity				
Lindgren et al., 2003 Sweden	CEA of cabergoline vs levodopa	Early stage PD patients.	Markov cohort simulation with transit probabilities derived from a trial	Time in HY stage, Mortality, Years of motor complications avoided, Direct medical costs	Five years (six-months with half-cycle correction)	TPP	Cabergoline was more expensive but more effective than levodopa	Yes
Smala et al., 2003 Germany	CEA of cabergoline vs levodopa	Early stage PD patients.	Markov cohort simulation with transit probabilities derived from a trial subset	Time in HY stage with/without motor complications, Proportion of patients without motor complications, Mortality, UPDRS score, Direct medical and indirect medical costs, Patient costs, Lost income, Informal care.	10 years (one-year)	Societal	Cabergoline was more expensive but more effective levodopa	Yes
Palmer et al., 2002 USA	CEA of entacapone+levodopa vs levodopa alone	PD patients with motor fluctuations	Markov cohort simulation with transit probabilities derived from trials	Time with ≤25% off-time/day, Mortality, QALYs, Direct and indirect health care costs Informal home care	Five years (Six-months with half cycle correction)	Societal and TPP	Entacapone was more expensive but more effective than levodopa	Yes
Linna et al., 2002 Finland	CEA of entacapone+levodopa vs levodopa alone	PD patients with motor fluctuations	Markov cohort simulation with transit probabilities	Time in modified HY stage, Mortality, QALYs,	Five years (Six-months with half cycle	Societal	Entacapone was cost saving compared to levodopa	Yes

			derived from 2 trials	Direct health care costs	correction)			
Nuijten et al., 2001 Netherlands	CEA of entacapone+levodopa vs levodopa alone	PD patients with motor fluctuations	Markov cohort simulation with transit probabilities derived from trials	Time with ≤25% off-time/day, Mortality, QALYs, Direct and indirect health care costs	Five years (six-months with half cycle correction.	Societal	Entacapone was cost saving	Yes
Shimbo et al., 2001 Japan	CEA of bromocriptine or pergolide compared to levodopa alone	Male PD patients treated with levodopa	Markov cohort simulation with transit probabilities derived from cohorts and clinical trials	Time in HY stage, Mortality, QALYs, drug costs, hospitalisation, informal home care costs	Ten years (one-year)	Societal	Both dopamine agonists dominated levodopa in late PD (HY stage 3+) but not early stage (HY 2)	None
Davey et al., 2001 Australia	CEA of pergolide vs bromocriptine	Early stage PD patients	Markov cohort simulation with transit probabilities based on cohort data adjusted by trial data	Time in early HY stages 1-3. Mortality, Direct medical costs. Nursing home costs	Ten years (Six-months).	TPP	Pergolide dominated bromocriptine with lower costs and longer stays in earlier HY stages	Yes
Hoerger et al., 1998, USA	CEA of pramipexole vs no drug or levodopa	Early PD, levodopa-naïve patients	Mathematical simulation with equations based on trial and survey data	UPDRS score, Mortality, QALYs Direct healthcare costs Nursing home placement Productivity	Lifetime (Not stated)	Societal	Pramipexole was more expensive but more effective than treatment without pramipexole	Yes
	CEA of pramipexole+ levodopa vs levodopa	Late PD, levodopa-treated patients						
Screening								
Van Laere et al., 2008 Belgium	CEA of SPECT imaging vs clinical evaluation	Uncertain PD diagnosis	Markov cohort simulation based on diagnostic accuracy	Adequately treated years (ATY), Mortality, Direct medical	Five years (Six-months)	TPP	SPECT was cost-effective achieving gains in ATY at low marginal cost	Yes

				costs				
Antonini et al., 2008 Italy	CEA of SPECT imaging vs clinical evaluation	Uncertain PD diagnosis	Markov cohort simulation based on diagnostic accuracy	Adequately treated years (ATY), Mortality, Direct medical costs	Five years (Six-months)	TPP	SPECT was cost saving	Yes
Dodel et al., 2003 Germany	CEA of SPECT imaging vs clinical exam	Uncertain PD diagnosis	Decision tree simulation based on diagnostic accuracy, prevalence data and expert opinion	Adequately treated months (ATM) Direct medical costs	12 months (Not stated)	TPP	Cost-effectiveness of SPECT conditional on willingness to pay to avoid inappropriate treatment	None
Surgery								
Hjelmgren et al., 2006 Sweden	CEA of a novel PD intervention vs standard drug therapy	PD patients with severe motor fluctuations	Markov cohort simulation based with time in each state based on regression of cohort data	Time in HY stage, Mortality, QALYs, Direct health care costs, Direct home care costs	25 years (Not stated)	Not stated	The novel procedure was cost saving over a 25 year time horizon.	None
Tomaszewski and Holloway, 2001 USA	CEA of deep brain stimulation (DBS) vs best medical management	Late stage PD patients with severe motor fluctuations	Markov model embedded in decision tree cohort simulation with event probabilities based on expert opinion and case series	Time in nursing home, UPDRS score, Mortality, QALYs, Direct medical costs, Nursing home care, Formal home care	Lifetime (one-year cycles with half-cycle correction)	Societal	DBS was more expensive but more effective than best medical care	None

* Manufacturer support

CEA = Cost-Effectiveness Analysis, TPP = Third Party Payer, HY = Hoehn Yahr (stage), CMA = Cost-Minimisation Analysis, UPDRS = Unified Parkinson's Disease Rating Scale, SPECT = Single Photon Emission Computed Tomography.

Table 2 Critical appraisal of the structural and data dimensions in the included models (summary) (work in progress/guide)

Intervention types/studies	Rationale for model structure		Disease states reflect disease progression?	Pre-model analysis			Uncertainty		Consistency	
	Evidence cited?	Reasonable?		Baseline progression	Treatment	Utilities	Methods	Structure	Internal	External
Medications										
Haycox et al	?/Yes	Yes	?/No	Yes	Yes	?	No	No	No	No
Hudry et al	?/Yes	/?Yes	?/No	Yes	Yes	?	No	No	No	No
Findley et al	?/Yes	Yes	?/Yes	?	?	Yes	Yes	No	No	Yes
Iskedjian and Einarson	?/Yes	?/No	?/No	?	?	N/A	No	No	No	No
Lindgren et al	No	No	?/Yes	Yes	Yes	N/A	Yes	No	No	No
Smala et al	?/Yes	?/No	?/Yes	?	?	N/A	Yes	No	No	No
Palmer et al	?/Yes	?	?/No	Yes	?	Yes	Yes	No	No	Yes
Linna et al	?/Yes	Yes	?/Yes	Yes	Yes	Yes	No	No	No	No
Nuijten et al	?/Yes	?/Yes	?/No	Yes	?	?	Yes	No	No	No
Shimbo et al	?/Yes	Yes	?/Yes	No	No	Yes	No	Yes	No	No
Davey et al	No	No	?/Yes	No	No	N/A	Yes	Yes	No	No
Hoerger et al	No	No	?/Yes	No	Yes	Yes	No	No	No	No
Screening										
Van Laere et al	Yes	Yes	N/A	N/A	N/A	N/A	No	No	No	Yes
Antonini et al	Yes	Yes	N/A	N/A	N/A	N/A	No	No	No	No
Dodel et al	Yes	Yes	N/A	N/A	N/A	N/A	No	No	No	No
Surgery										
Hjelmgren et al	?/Yes	?/Yes	?/Yes	Yes	Yes	Yes	Yes	Yes	No	No
Tomaszewski and Holloway	No	No	?/Yes	?	?	?	Yes	Yes	No	No

NA = not applicable; Reviewer opinion Yes = criterion substantially met; No = criterion not met or absent; ? = unclear whether criterion was met; ?/Yes = criterion partially met; ?/No = criterion present but problematic/debatable

Generally each criterion needs to be clearly stated, consistent with other criteria, and justified as appropriate. Although inherently a matter of subjective opinion, the above table highlights broad areas of strength and weakness in the modelling techniques which are considered in more detail in the narrative text.