Challenges in detecting disease modification in Parkinson's disease clinical trials

Dr Dilan Athauda, Sobell Department of Motor Neuroscience, UCL Institute of Neurology & The

National Hospital for Neurology and Neurosurgery, Queen Square, London, WC1N 3BG

Corresponding author: Athauda, D.S. (d.athauda@ucl.ac.uk)

Telephone: +44 203 448 8739, Fax: +44 203 448 8642

Dr Thomas Foltynie, Sobell Department of Motor Neuroscience, UCL Institute of Neurology & The

National Hospital for Neurology and Neurosurgery, Queen Square, London, WC1N 3BG

Email address: t.foltynie@ucl.ac.uk

Telephone: +44 203 448 8726, Fax: +44 203 448 8642

Abstract

Despite the wealth of encouraging data from numerous compounds that demonstrate

"neuroprotection" in pre-clinical studies of Parkinson's disease, and despite numerous clinical trials,

to date, no intervention has been demonstrated to able to modify the course of disease progression.

While this "failure to translate" is likely due to numerous factors including our incomplete

understanding of the pathogenic mechanisms underlying PD together with excessive reliance on

data from the toxin-based animal models of PD, here we will discuss the "structural issues"

pertaining to inadequate clinical trial design, selection of inappropriate endpoints and poor patient

selection which are often not addressed following failed disease modification trials. Future

directions to overcome these challenges such as reducing the heterogeneity of patient cohorts,

identifying and utilising a pre-diagnostic population, embracing a personalised medicine approach and utilising novel trial designs may be required to ultimately fulfil the goal of conclusively demonstrating evidence of disease modification.

Despite the development of several effective pharmacological and surgical therapies for the treatment of Parkinson's disease (PD), the ultimate goal of slowing of disease progression has not been met. Over time motor symptoms of tremor, rigidity and bradykinesia are invariably complicated by continued degeneration of non-dopaminergic systems leading to mood and behavioural issues, dementia and postural instability leading to worsening functional disability and self-care. Pre-clinical studies have focused on halting or preventing this dopaminergic cell loss and numerous "neuroprotective" compounds have been demonstrated with the hope that these properties will translate to human subjects and slow the progression of symptoms and modify the expected disease course. However, despite the promise of many therapies in pre-clinical trials, no disease modifying trials conducted to date have conclusively demonstrated evidence of disease modification.

This general "failure to translate" is likely related to numerous factors including our incomplete understanding of the pathogenic mechanisms underlying PD together with excessive reliance on data from toxin-based animal models of PD to judge which agents to take to double blind clinical trial evaluation. Detailed discussion of these factors is outside the scope of this article and will not be discussed here (see [1] for review).

Inherent errors in data collection when conducting clinical trials are mostly unavoidable and often result in diluted effects, the need for larger sample sizes and increased costs. However, "structural errors" pertaining to inadequate trial design, selection of inappropriate endpoints, poor patient selection, or mishandling of missing data, can be minimised and increase the chance of detecting true disease modifying effects [2,3]. This review will consider methodological issues relating to detecting efficacy in disease modification phase III trials and suggest future directions that are being explored that may improve the ability to detect any signal of effect. Earlier stages in drug development and clinical trials present their own specific challenges and these will not be addressed (see [4,5] for review).

Target population for clinical trials- embracing personalised medicine

PD patients exhibit remarkable heterogeneity - studies suggest that different pathophysiological mechanisms relate to different clinical subtypes and this disease heterogeneity can influence disease progression [6,7]. Therefore it may seem somewhat counter-intuitive to gather such a heterogeneous group of patients and expect a uniform response to a particular intervention. Nevertheless, in all disease modifying trials to date this has been the general approach.

Common methods of refining patient selection are to recruit patients only at a specific stage of disease, although this is by no means a way of predicting subsequent rates of disease progression. Nevertheless, in considering patient recruitment to trials of potential disease modifying interventions, patients can be divided into two categories – early untreated "de-novo" symptomatic patients and those patients on stable dopaminergic treatment that experience disability – both of which have merits and disadvantages for inclusion into disease modifying trials (Table 1).

Another factor when selecting patients is that motor progression in PD is not linear and is thought to occur in phases. Factors such as age of onset, predominant motor phenotype and presence of non-motor symptoms can influence disease severity at presentation, rate of disease progression, and most importantly, response to treatment [8], suggesting the existence of various sub-types of PD. These baseline differences are often not always appreciated when incorporating data into the primary outcome.

Future directions - improving patient selection and stratification for clinical trials-

Reducing heterogeneity of cohorts

Methods to reduce the heterogeneity of patient populations and enable the prediction of rates of disease progression may reduce inherent variability that currently exists between individual participants in a trial and thus increase its statistical power, in turn enabling more efficient and cost-effective collection of data and increase the likelihood of detecting a signal of effect.

Although there are ongoing efforts to identify PD subtypes, well-defined criteria have not yet been widely accepted, but by using clinical, biochemical or genetic markers (or a combination) researchers may be able to improve patient stratification for inclusion of more homogeneous cohorts into future trials (Figure 1).

Basic clinical markers (i.e. motor phenotype) can be potentially useful to predict disease progression. Analysis of data from the Parkinson's Progression Markers Initiative (PPMI) database suggests that older age of onset is associated with a more severe motor and non-motor phenotype, greater dopaminergic dysfunction on DaTSCAN, and reduction of CSF alpha synuclein and total tau [9]. Similarly tremor-dominant PD is associated with slower disease progression and less-severe cognitive impairment than akinetic-rigid PD [10], while a recent study demonstrated that the presence of mild cognitive impairment, orthostatic hypotension and REM Sleep behaviour disorder at baseline is associated with a more malignant disease course [11]. Building on this, a recent model has been validated that allows individual patient prognostication. Based on 3 predictor variables, namely higher patient age, higher Unified Parkinson's Disease Rating Scale motor examination axial score, and a lower verbal fluency score for animals at baseline, patients can be reliably predicted to be more likely to develop non-levodopa responsive symptoms of dementia and postural instability at 5 years from diagnosis [12]. These simple variables could be used to create a more homogenous cohort with a predictable disease course.

Clinical markers may however be unstable in early disease and may also be substantially influenced by symptomatic therapy [13]. Recent studies suggest that biochemical markers may be able to more precisely identify molecular subtypes that have differential responses to treatment. A recent study suggested that patients with low levels of Aβ1-42 or highest t-tau/Aβ1-42 had more severe non motor dysfunction, whereas lower alpha-synuclein levels were associated with worse cognitive performance [14] and therefore may potentially be useful to predict patient subgroups likely to have accelerated disease progression. Similarly, the presence of high levels of pro-inflammatory immune

markers in patient serum is associated with more rapid motor progression and lower MMSE scores [15].

Precision / personalised medicine

Additionally, identification of genes involved in familial PD has suggested that distinct pathophysiology could also underlie different forms of PD [16] and also affect clinical phenotype and motor progression [17]. This may allow trials to offer "personalised" interventions that specifically target an aspect of PD disease pathophysiology known to be abnormal in individual patients and recruit patients accordingly, increasing the chance of influencing disease progression. A recent study using neural cells generated from induced pluripotent stem cells derived from PD patients and pre-symptomatic individuals carrying mutations in the PINK1 (PTEN-induced putative kinase 1) and LRRK2 (leucine-rich repeat kinase 2) genes [18] analysed the subsequent cellular responses when treated with potential neuroprotective agents coenzyme Q10, rapamycin and LRRK2 inhibitors. The study indicated that treatments had varying levels of effectiveness, depending on the genetic mutation involved; for example rapamycin reduced oxidative stress and mitochondrial stress in LRRK2 mutant neural cells but not in PINK1 patient neural cells. This technology could be used to predict drug efficacy, and identify potential drug-responsive cohorts for selection in clinical trials.

A precision or personalised medicine approach in combination with genetic and epigenetic testing is commonplace in patient selection for cancer clinical trials and has yielded numerous successes. Patients with breast cancer that overexpress human epidermal growth factor receptor 2 (HER2) have been found to not respond to standard therapy but by studying a homogenous population all with this particular mutation, it has been found that trastuzumab, a monoclonal antibody can be effective in this particular subgroup [19].

Similarly, this approach is starting to be utilised in PD clinical trials. Mutations in the gene encoding glucocerebrosidase (GBA) constitute the most frequent risk factor for PD in the general population,

and the presence of a mutation in GBA confers a worse prognosis to PD patients. Studies show that GBA mutations are associated with reduced glucosylceramidase activity and oxidative stress while simultaneously it has been demonstrated that ambroxol, a secretolytic agent licensed for the treatment of respiratory diseases, can increase glucosylceramidase activity and reduce alpha-synuclein levels in pre-clinical models [20]. A small clinical trial specifically recruiting patients with mutations in the GBA gene is currently underway to assess the safety and dosing range of ambroxol and represents a novel method of selective recruitment based on possible mechanism of action. However it remains to be seen whether highly targeted drugs or drugs that target different systems will be most successful.

As with all these possible methods, uncertainty remains regarding the extrapolation of such results in a wider heterogeneous population and so further studies may have to be explored to assess the response in a more heterogeneous population.

Maximising neuroprotective potential

Trials that select patients with "early" or untreated disease are commonly recruited to maximise any possible neuroprotective effects however, even in this population neurodegeneration may be substantially advanced by the time of diagnosis. A study quantifying the extent of nigrostriatal degeneration at varying stages of PD 1 to 27 years post diagnosis demonstrated that tyrosine hydroxylase-positive neurons in the fibres of the dorsal putamen are already reduced at diagnosis, and are virtually absent as early as 4 years after diagnosis [21]. Therefore identifying and recruiting patients at the very earliest stages or even prior to neurodegeneration may represent a better strategy. Studies suggest the onset of PD can begin up to 20 years prior to appearance of classical motor symptoms while data from imaging studies utilising PET and SPECT suggest nigrostriatal degeneration can be detected approximately 5-10 years before the appearance of classical motor symptoms [22,23]. This pre-diagnostic period includes completely asymptomatic individuals at risk of developing PD ("pre-clinical") as well as those patients exhibiting mild motor and non-motor

symptoms although insufficient for a diagnosis of PD ("pro-dromal"). The pre-diagnostic population

represents an appealing group for inclusion as any disease modifying therapy could maximise

benefits by potentially delaying or completely preventing the onset of overt disease, by preserving

greater number of neurons and preventing secondary self-propagating degenerative processes.

However, accurately identifying this pre-diagnostic population is very difficult. Although

environmental factors such as exposure to pesticides and heavy metals such as manganese, and

lifestyle factors such as avoidance of smoking, caffeine and alcohol have all been associated with an

increased risk of PD [24], the low predictive values and relevance of these factors hinder their utility

in identifying patients in longitudinal studies. Even in patients with a single identifiable strong risk

factor for PD such as idiopathic RBD, being a LRRK2 carrier or possessing a GBA mutation, there is

wide variability of conversion to manifest disease[25,26].

Prodromal PD describes the stage in the natural history of PD wherein patients exhibit subtle motor

and non-motor symptoms and signs but which are insufficient to diagnose PD.   Patients in the

prodromal stage of PD offer similar advantages as the "at risk" population to researchers of disease

modifying trials, however similar issues with the accurate identification of these patients also exist.

Recently, a task force from the MDS developed research diagnostic criteria for the prodromal phase

of PD based on the presence or absence of certain risk factors such as sex, RBD, depression,

abnormal substantia nigra echogenicitiy etc. These criteria were recently successfully applied

retrospectively to a longitudinal study cohort of elderly patients with a sensitivity of 'probable

prodromal PD' status for incident PD of 54.6% (95% confidence interval, 28.0–78.8) and specificity of

99.2% (97.8–99.8), indicating that it may be useful in defining target populations for disease

modifying trials.

However, limited accessibility to specialised functional imaging and diagnostic techniques such as

polysomnography may limit identification of this population, though if patients could be reliably

identified, there remain new issues with regards to selecting an appropriate endpoint.

The first appearance of motor dysfunction as detected by the MDS-UPDRS may indicate appearance of disease onset (and has been used as an endpoint in premanifest Huntington's disease trials), however mild elevation in MDS-UPDRS Part III scores (2-10) can occur in the general population and increases with age [27], making interpretation difficult. Conversion to overt PD could in theory be used as an endpoint. Given there is a only low risk of incident PD (1.67% of individuals over 50yrs after 5 years in a recent observational study [28]), trials would have to recruit extremely large numbers of healthy individuals and have unfeasibly long follow up periods to be able to detect even a large risk reduction from any disease modifying therapy [26]. In contrast, enriching patient cohorts with individuals with a high combination of risk factors that influence disease progression (such as patient with idiopathic RBD, hyposmia, and with abnormal DAT-SPECT imaging) may increase the incidence of subsequent PD diagnosis and hence reduce the number of participants required to adequately power a trial of a potential disease modifying therapy [26]. Though currently this would still likely require large numbers of patients followed up for a number of years to detect a small reduction in PD risk. Furthermore, there are inevitable ethical concerns in informing individuals that they are at high risk of developing PD, especially while no disease modifying therapies currently exist, and strategies to deal with this might include blinded inclusion of both high risk and a subset of low risk patients.

A significant issue with utilising a pre-diagnostic population in an interventional trial is that exposing asymptomatic or minimally affected individuals to novel therapies with potential adverse effects may induce high patient drop out. Additionally, there remains the inherent difficulty of assessing the long-term safety of these new treatments, which may target molecular mechanisms involved in other diseases. For example, efforts to target the toxic effects of alpha-synuclein aggregation as a disease modifying strategy using passive and active immunisation with monoclonal antibodies against alpha-synuclein have been found safe in phase 1 trials [29]. However there are theoretical concerns that some species of alpha-synuclein may be protective [30] and the long term consequences of lowering alpha-synuclein may have unforeseen outcomes [29]. Alternatively using

drugs already in clinical use with well-known safety and pharmacokinetic profiles and "re-positioning" them for use in PD may allay some of the short term safety concerns but often long term safety data regarding their use in a novel population (e.g. PD patients) is lacking.

Clinical trial design

Although previous trials have putatively reported potential disease modifying effects [31,32],controversy remains due to significant difficulties in being able to distinguish between any symptomatic effects of the therapy and any true disease modifying effects. There are a number of approaches to trial design that have attempted to disentangle these mechanisms, each with their own advantages and disadvantages.

Washout studies involve a two period design that includes a "washout" phase at the end of the study (whereby the intervention is withdrawn) prior to the final assessment of the patient and evaluates the change in Parkinson's disease rating scale scores. Any benefit seen in the active arm that persists over the washout period may indicate evidence of disease modification. This approach has been previously used to assess deprenyl (Selegiline) vs placebo [33] and levodopa vs placebo (ELLDOPA)[32] on disease progression in PD. However, although both trials seemed to indicate persistent beneficial effects after the washout period, many have questioned the validity of the results and subsequently regulatory approval of these drugs as disease modifiers has not been accepted. An inherent issue with this design is that is if the washout period is too short, uncertainty remains regarding whether any of the persistent beneficial effects seen could be due to unanticipated long-lasting symptomatic effects. Washout periods are typically calculated as some multiple of the half-life of the pharmaceutical product within the population of interest; however this may not always be the most optimum method. The ELLDOPA trial utilised a washout period of 14 days, initially thought to be sufficient based on the half-life of levodopa (90minutes), however, it

is now known that symptomatic effects of levodopa can still be detectable up to 32 days after withdrawal [34]. Additionally, differential bias may occur in the evaluation of those declining more rapidly in the washout phase. Alternatively, using a long washout period of several months to ensure loss of symptomatic effects could result in a confound by the progressive nature of PD, induce significant lessebo effects [35] and risk high levels of patient dropout in untreated patients, and may have ethical considerations.

Delayed-start designs are a promising alternative to wash out studies and comprise a similar two period design. The first period involves patients randomised to either placebo or active intervention followed by the second period whereby both groups receive the active intervention and are followed up for a fixed time period. Exposing patients to placebo at the start of the trial combined with all participants eventually being exposed to active intervention may improve subject recruitment and retainment and may minimise any lessebo effects in the second period. Although there is no agreed statistical method for analysis and interpretation of the data for delayed start designs, the most common method used to identify evidence of disease modification relies on analysis of 3 components in sequence – the superiority of treatment over placebo during the first period; the non inferiority of early vs. delayed start (i.e. that slopes during the second phase comparing both groups on active treatment do not converge); and that there remains a significant difference between the two groups at the end of the study period[36]. Analysis of the data requires complex statistical analysis and multiple hypotheses in a sequential manner, but this design has been used successfully in a number of trials assessing the disease modifying effects of rasagiline (TEMPO[37], ADAGIO[31]), pramipexole (PROUD[38]) and deferiprone (FAIRPARK[39]) and remains of great interest to researchers [40]. However, there are a number of problems with this design. The initial period has to long enough for any neuroprotective effects to be detected, typically estimated at 6-12 months (which also raises ethical concerns of (potentially drug-naïve) patients remaining on placebo for a long period and risks high patient drop out, leading to an over representation of individuals with minimally progressive symptoms). Delayed start designs must be powered to detect small changes

in UPDRS, and while the relevance of such small changes is uncertain, they also generally need longer follow up to ascertain whether any differences is sustained. For example, the minimal change in total UPDRS score thought to represent a clinically meaningful improvement for patients is around 5-8 points [41–43] and the ADAGIO[31] and TEMPO[37] trials demonstrated a difference of around 2 UPDRS points between delayed start and early start groups. In the FAIRPARK trial, small differences between the delayed start and early start groups waned after a longer follow up, calling into question the relevance of such small differences [39]. Also, although the design of the trial is "double-blind", as the second period involves all participants on the active intervention, it is questionable whether the blind can be fully preserved.

An alternatively used approach are trials not focused on differentiating symptomatic from disease modifying changes but rather changes in long term disability ("milestones") that are clinically meaningful. Longitudinal study "single-period" designs use composite global measures assessing a range of motor and non-motor symptoms for a multidimensional assessment of disease progression over a long time period, typically a number of years, are considered most effective in assessment of disease-modifying therapies. Patients could be recruited all varying stages of the disease, however such trials require considerable investment and extensive collaboration between pharma companies, research institutes and clinical centres to manage the often substantial costs.

Futility trials can be used to screen out therapies unlikely to prove efficacious in larger phase III studies and use a short term outcome to compare the treatment group against a historical control (with therapies considered futile if alpha is greater than 0.10 or 0.15 (a P value less stringent than typically accepted to reduce sample size and false negatives), thus requiring a fraction of the patients and resources used in more conventional studies [44]. However, a major limitation is that the choice of data used to compare futility must be carefully considered as historical controls have been shown to behave differently from modern placebo groups [45] and so more recent trials have

utilised a concurrent placebo group to overcome this confound [46] – however this may eliminate many of the futility designs' original advantages.

Adaptive designs utilise interim analysis to modify the course of a study without affecting the integrity of the trial. Newly available, high quality evidence can be incorporated into the study design allowing for alteration of sample size based on observed variability, modifying ineffective doses, moving patients among various trial phases, or simply ending a trial early due to failure of efficacy (such as in the case of NET-PD and PRECEPT studies [47,48]). Although these designs can improve efficiency and reduce drug development time, they may introduce bias, require complex operational support from statisticians and safety monitoring boards in order to make real time decisions based on current data [2,49]. Additionally, in both futility and adaptive designs, characterising a response as "positive" or "negative" in a relatively short time period in a such a slowly progressive condition as PD can be challenging.

Future directions – improving data collection and trial design

Selecting the appropriate trial design, reducing the inherent variability of data collection and utilising novel methods will be essential for maximising the potential to detect any positive signal of effect and may reduce the often prohibitive costs of clinical trials of putative disease modifying agents (Figure 2).

Trial design and disease modelling

It must be emphasised that properly randomised, double-blind, placebo controlled trials currently remain the most accepted standard for assessing disease modification in clinical trials. However utilising "pragmatic" trial designs may overcome some of the economic and time-consuming limitations these trials have.

An example of using a less conventional approach was recently utilised to generate preliminary efficacy data regarding exenatide treatment on disease progression in PD. In the absence of placebo

control, investigators used a parallel group of patients with moderate stage PD on stable drug treatment that were matched to the treatment group, but did not receive injectable therapy – allowing a comparison of the novel treatment with best available therapy [50]. Patients were followed for 12 months followed by a final assessment 2 months after exenatide withdrawal (allowing for any symptomatic effects to dissipate). Additionally a longer washout period of 12 months followed the end of the active intervention theoretically enabling any long lasting symptomatic or placebo effects to dissipate and was able to be tolerated as patients were on stable medication regimes [51]. Although not as robust as a double-blind, placebo controlled trial, this data was used to justify further funding for a larger, Phase II study and has suggested that clinically informative data might successfully be obtained from patients with PD using standard clinical measures, without biomarkers and without a placebo-controlled arm, in a relatively inexpensive way [52]. However, as evidenced by previous studies demonstrating the importance of the placebo effect in PD patients and particular in clinical trials [35,53,54], there remains understandable uncertainty regarding the use of data from studies that do not utilise a placebo-controlled arm.

Novel techniques such as clinical trial simulation and disease modelling have advanced in recent years and could be used to minimise or avoid extrinsic errors in trial design and aid researchers to select an appropriate design to utilise. The results of a double-blind, placebo-controlled trial of intraputamenal infusion of GDNF as a treatment for PD excluded the possibility of a clinical effect [55], despite encouraging results from previous open label trials. However, although reasons such as methodical differences and placebo effect were proposed to explain these discrepancies, further analysis of the statistical method suggested the trial was underpowered and therefore incapable of drawing conclusions from the seemingly negative results[56]. Disease modelling and simulation can now be used to facilitate design of clinical trials and calculate the probability of a specific outcome [57]. Using these techniques, a study suggested washout designs may be superior to delayed start designs to detect evidence of disease modification (assuming a washout period of at least 3 months)

[58] though as previously noted, long washout periods can themselves be problematic for untreated patients.

Additionally, using disease progression data from natural history studies and pharmacodynamic and pharmacokinetic models, data can be integrated to simulate different trial scenarios for predictive purposes. A clinical trial stimulation model for AD has been recently been developed and accepted by the FDA as a quantitative tool that could be used to calculate variables such as sample size, bias and power between varying study designs, and if successful, further models regarding PD may be developed.

Reducing inherent errors / variability of data

Dealing with missing data is a problem inherent to all trials but varying methods can significantly affect results. Recently new statistical methods to overcome methodological issues in the delayed start design such as discontinuation rates and missing data to test non inferiority have been proposed and successfully applied to Phase 3 AD clinical trial data, suggesting it can control Type 1 errors and provide robust evidence of disease modifying effects, and this could be utilised in future PD trials [36].

In longitudinal studies, patient dropout is not uncommon but the resulting missing data can be a major complicating factor in the analysis, and may have substantial consequences for the interpretation of the results [59]. As well as efforts to minimise patient withdrawal, studies have shown that Mixed-Effect Model Repeated Measure (MMRM) analysis appears to be a superior approach in controlling Type I error rates and minimizing biases, as compared to the more traditional "last observation carried forward" (LOCF) ANCOVA analysis [60], and with relevance to disease modifying trials, regulatory agencies have recommended the MMRM approach.

Selecting suitable endpoints

Demonstrating evidence of neuroprotection in pre-clinical studies is relatively straightforward, however quantifying neuronal cell death is not a feasible end point in clinical trials and therefore demonstrating that a therapy has delayed clinical disease progression or postponed disability is a pragmatic and acceptable alternative. In view of this, alternative surrogate endpoints as markers of disease progression have been utilised in previous clinical trials, though all have their own advantages and disadvantages.

Time to onset of event

The time taken for a drug-naïve patient to require dopaminergic or symptomatic therapy is often used in clinical trials as evidence of disease progression. Although this can be a well-defined, measurable endpoint [61] that is not confounded by symptomatic effects from other drugs , a notable limitation is that in clinical practice, the decision to start symptomatic therapy can be influenced by a number of factors that depend on the patient, physician and care centre. A study showed that almost 50% of patients required dopaminergic therapy before the end of 12 months [62], and this was influenced by factors such as level of patient education and emotional state. The time to require therapy, therefore, does not necessarily correlate with disease stage or rate of progression, and it is often impossible to differentiate whether a delayed need for symptomatic treatment reflects symptomatic or neuroprotective effects of the studied drug [61,63].

The onset of dementia, postural instability and need for residential care are well recognised late stage clinical milestones that occur at a similar time from death independent of age of onset [64] and as such have been suggested as suitable endpoints to study. However, the interval between diagnosis and development of these milestones is extremely heterogeneous; 10-year follow up data from the CamPaIGN study showed that although most patients at 10 years had either developed dementia or postural instability, 25% of patients still had preserved mobility and cognition[65]. Practically speaking this would mean conducting trials with extremely long follow up or selectively recruiting patients with advanced disease (with their inherent disadvantages as noted above).

However, using data from natural history studies patients could be tracked against their individual expected disease trajectory which could allow the detection of true disease modification and potentially remove the need for control arms [66].

Changes in disease rating scales over time

Measuring the progression of clinical features over time is another commonly used method of assessing therapeutic efficacy and these can be divided into those that measure impairment (e.g. MDS-UPDRS) , disability (e.g. the Schwab and England Activities of Daily Living (ADL) Scale (S&E) or quality of life (e.g. the 39-item Parkinson's disease questionnaire ([PDQ-39]) [67]. Although in essence these assess the symptomatic effects at a single time point, repeated measurements combined with methods to delineate between symptomatic or disease modifying effects (delayed start design, washout etc.) or comparison to control groups can be used to demonstrate delay in disease progression. Additionally, data collected from placebo groups can be a useful resource for future trial planning. The MDS-UPDRS is the most widely used and accepted endpoint, is clinically relevant to patients, assesses motor and non-motor symptoms, has been validated and calibrated against the older UPDRS, has objective instructions for use and uses a certification process to reduce inter-rater variability [68,69]. Despite this, use of the MDS-UPDRS as a primary end point has a number of issues. In patients on stable therapy, symptomatic medication influences the assessment of the UPDRS Part III (the motor subscale). Furthermore, being able to define a minimally clinically important change in UPDRS scores that is relevant to patients is critical in order to interpret the relevance of a statistical change and to aid calculation of sample size for future trials. The annual rate of change of the older UPDRS total score estimated from older interventional trials is around 8-10 points per year and a minimally relevant clinical change is around 5-8 points, but equivalent rate of change for the MDS-UPDRS has not yet been clearly established. Additionally, a recent study has shown that the rate of progression of UPDRS scores is not linear, but rather occurs in three distinct phases following diagnosis: improvement, stability, and steady decline [8], and this rate of change is

also influenced by variables such as age at diagnosis, baseline motor score and cognitive status, which are often not controlled for in the analysis. Also, some symptoms within the scale do not progress evenly - data from longitudinal studies have shown that although average UPDRS motor scores increased by 1.2 points/year, tremor scores did not change over time [70] – perhaps underlining the importance of stratifying patients according to phenotype.

To increase the chance of detecting a signal of effect, longitudinal assessment of the MDS-UPDRS scores are also often supported by measurement of a wide range of secondary end points such as scales to assess effects on quality of life, cognition, mood, dyskinesia, gait, and non motor systems. Together these can allow detection of effects on non-dopaminergic systems also, as well as motor and non-motor symptoms. The rationale is that any potentially positive effects of disease modifying therapy would be more convincing if they are also shown to have benefit across a range of dysfunctional neurotransmitter systems. While these questionnaires are thought to provide a global indicator of function, further work is needed to validate their usage as measure of disease progression. Patient diaries as end points in clinical trials have been associated with reduced compliance, recall bias and diary fatigue, particularly in patients with cognitive dysfunction and depression[71], and often do not correlate with quantitative objective assessments [72,73]. Stage of disease can also influence the usefulness of the questionnaires. A study suggested that UPDRS scores correlates well with disease progression, but questionnaires assessing overall disability such as the PDQ-39 and Total Functional Capacity scale were not sensitive to change in early disease [47], whereas in moderate stage disease, the PDQ-39 summary index was found to be most sensitive to change [74].

Change in imaging markers

As a result of the limitations of clinical evaluations, many trials have utilised surrogate functional imaging markers of the nigrostriatal dopaminergic system as a primary or secondary endpoint to act as a measure of disease progression. Previous studies have measured dopamine terminal function

by 18F-Dopa PET, or dopamine transporter function, using PET, using tropane-based tracers, or single-photon emission computed tomography (SPECT), with a variety of ligands [75,76]. These markers have the advantage of being able to detect changes in patients with early disease and are more objective than clinical assessments. However, modalities such as PET imaging are often prohibitively expensive, with availability and standization across centres being a further issue. Also, there remain a lot of issues regarding validation and qualification of these methods. For example, in the ELLDOPA trial, using B-CIT SPECT as a surrogate imaging marker, there was faster decline in striatal dopamine transporter in patients on L-dopa despite a clear clinical improvement, and other trials have demonstrated similar discordant results between imaging and clinical progression [55,77–79]. Confounding effects of symptomatic drugs on proteins interacting with radioligands has since been postulated as a possible explanation for these effects and has highlighted that there may still some way to go before these techniques become accepted as validated biomarkers in efficacy trials by regulatory agencies.

Future directions – reducing variability and improving quality of data

Research into biomarkers

Identifying PD-specific, sensitive, validated quantifiable biomarkers are alternatives to subjective clinical rating scales and would greatly enhance techniques to measure disease progression and therapeutic interventions. Ongoing research into using patient tissue and fluid biomarkers is advancing including CSF analysis (see [80] for review), biochemical tests (see [81] for review) and neurophysiological tests (olfaction, motor tests) to investigate measures that most accurately measure disease progression but further work is need to validate and qualify their use before they are accepted by regulatory agencies.

There are similar advances in imaging biomarkers - new techniques using 7T MRI, novel MRI sequences (utilising a combination of diffusion tensor imaging, T2*-weighted and whole-brain T1-

weighted techniques) , novel PET ligands or transcranial ultrasound have shown promise in early identification of patients prior to a clinical diagnosis of PD, though as markers of disease progression, results are less convincing. Quantification of abnormal deposition of alpha synclein may be a more direct measure of disease progression in PD and a recent study has demonstrated that a novel ligand can bind alpha-synuclein fibrils and may be able to be used to monitor disease progression [82]. Other non-dopaminergic targets may also be relevant - a recent study using 3T MRI scan demonstrated that cortical thinning in the temporo-parietal regions could be used as a biomarker of mild cognitive impairment in patients with PD progressing towards dementia [15,83].

Minimising variability / error in assessing endpoints

The lack of reliable CSF, blood or imaging biomarkers, means that clinical trials currently have to rely on clinical endpoints or symptomatic markers to assess disease progression. These measures are often insensitive, rely on episodic monitoring which only provides a "snap-shot" of the patient, can be influenced by natural disease fluctuation, are subjective and can be prone to inter- and intra-rater variability[84], and generally require the patient to attend the research centre to collect the data required, incurring expense and burden to the patient, which negatively impacts on patient recruitment [85]. Efforts to reduce variability of assessments and the "noise" of the effects of symptomatic medication on motor assessment scales include assessing patients after a period of overnight drug withdrawal (i.e. in the OFF medication state) though this may impact recruitment as some patients are unable to tolerate this. Similarly, in order to minimise rater variability, some studies advocate video recording assessments of MDS-UPDRS motor scale to act a "quality-controls" (though this does not affect patient expectation and the size of the placebo effect) and assessing the inter-rater reliability at the beginning of each clinical trial and select only those raters who have achieved good kappa values [86].

The evolution of mobile and online technology means it is now possible to be able to remotely collect objective data about disease progression from patients in a continuous and quantitative

fashion. Using wearable devices, various studies have shown that is possible to collect objective data regarding bradykinesia, tremor, dyskinesia, walking, sleep and speech that correlate to severity of disease, response to medication and track disease progression [87–90]. A recent study also showed that an intelligent closed-loop system integrating a wide range of wearable sensors, allowed physicians to remotely monitor the condition changes of the patients, adjust medication schedules, and individualize therapy in real-life conditions [91].

Many of these systems also allow home or continuous monitoring, allowing increasing detection of events currently undetectable, being able to capture and incorporate motor fluctuations [88], and the need to avoid episodic assessments in the research centre. Additionally, virtual visits using telemedicine can decrease burden on patients, improve patient recruitment; reduce variability in assessments and chances for missing data. Studies have shown that this technique is feasible and can offer similar benefits to that of in-person evaluations [92,93]. Although many of these devices compare favourably with the "gold standard" of assessments of motor function, the MDS-UPDRS, and can provide greater test-retest reliability and sensitivity to change than conventional clinical ratings [94], there are current limitations to such approaches. Some aspects of assessment such as rigidity are extremely difficult to assess remotely and these devices remain to be validated in large population of patients with PD.

While most endpoints aim to measure the change in clinical rating scales over time, a recent study has suggested a novel composite binary endpoint measuring mortality and dependency [67]. The novel measure was found to be clinically relevant, sensitive to disease progression from patient's already on stable symptomatic therapy (minimising patient drop out) and requires smaller sample sizes enabling trials to be more cost effective, though further validation in larger cohorts is required.

Yet despite all these refinements to optimise endpoints, there remains the question of whether positive results demonstrated in a given population during a (relatively) short duration disease modification trial can be extrapolated throughout the course of the disease in other populations.

The duration of many disease modification trials is determined by a number of factors but must be sufficiently long enough to allow the adequate detection of the expected treatment effect in the sample size being studied while also taking into consideration the many technical and economic factors as outlined above. As a result, any observed changes in disability are often small and the clinical relevance in wider populations and at varying stages of the disease is uncertain.

Adding the proposed experimental intervention to existing stable symptomatic therapies in a large, placebo controlled, longitudinal "one-period" trial, using clinical milestones and global measures of patient quality of life as endpoints would be one pragmatic method of overcoming these limitations [3] but incur many of the same issues as stated above and require considerable investment and resources.

Conclusions

Despite the continued failure to date of finding any disease modifying therapy in PD, encouragement can be sought from parallel work in the field of research into multiple sclerosis. MS, like PD, is a heterogeneous, (ultimately) neurodegenerative disease with marked variability in clinical phenotype as well as radiological pathological and genetic phenotypes. Yet despite such differences remarkable progress has been made in the development of drugs that can alter the trajectory of some forms of MS. However developing and utilising a study design that is able to detect evidence of delayed disease progression in PD is challenging. Although double-blind, placebo controlled trials remain the gold standard of trial design, more pragmatic designs have demonstrated that useful data can be gathered quickly and cost-effectively, paving the way for more robust studies. Reducing variability in data collection may be possible with advances in technology enabling efficient, objective and continuous data collection describing a wide range of symptoms. Additionally improved , and where necessary targeted patient selection may be crucial for future trials, with consideration that cohorts may be better focused on participants most likely to show a measurable change over the course of study. In parallel, a variety of outcome measures that reflect the diversity of range of PD phenotypes

are needed that are sensitive to the intervention and vary over time. Similarly, the establishment of valid biomarkers will alter how clinical trials measure disease progression and ongoing longitudinal studies may enable early identification of patients who would be most likely to gain from any neuroprotective therapies.

Table 1 - Patient suitability for disease modifying clinical trials

De novo untreated patients	Symptomatic stable treated patients
Advantages	**Advantages**
- Considered "early" disease – more salvageable neuronal population (30-50% of neurons lost)[1,2] - No cofounding effects of symptomatic medication - Minimal clinical fluctuation	- Easy to identify and recruit into trials - Clinical changes in disability more likely to be detected rating scales - Less likely to include recruiting patients without idiopathic PD
Disadvantages	**Disadvantages**
- Fewer in number, recruitment is slow, incurring greater expense - Patients with greater disability at baseline will have shorter "untreated" tolerable window - narrowing chance to detect disease modifying effects - Maintaining patients with mounting disability on placebo may be unethical - Increasing disability in some patients may induce high dropout rate – leading to over representation of patients with slower disease course - Drop out may lead to difficulties with data analysis - Risk of including patients with atypical forms of parkinsonism (and alternative pathologies) - 10% of patients clinically diagnosed with early PD have SWEDD who majority do not go to develop PD [3,4] - Clinical rating scales at early stage disease may not be sensitive to detect changes in disability and create floor effects - Heterogeneous clinical progression	- Symptoms influenced by effects of medication - Clinical fluctuation and variable response to medication make longitudinal assessments difficult "Ceiling" effects of clinical rating scales such as UPDRS make changes in advanced disability difficult - In trials with placebo arm, may exhibit increased magnitude of placebo effect - Advanced neurodegeneration may mean that any neuroprotective therapies may be "too little, too late" - Greater number of co-morbidities in advanced population may preclude lengthy longitudinal assessments [5] - Heterogeneous clinical progression

Figure 1 - Methods to stratify patients prior to selection in Phase 3 disease modification trials to reduce heterogeneity.
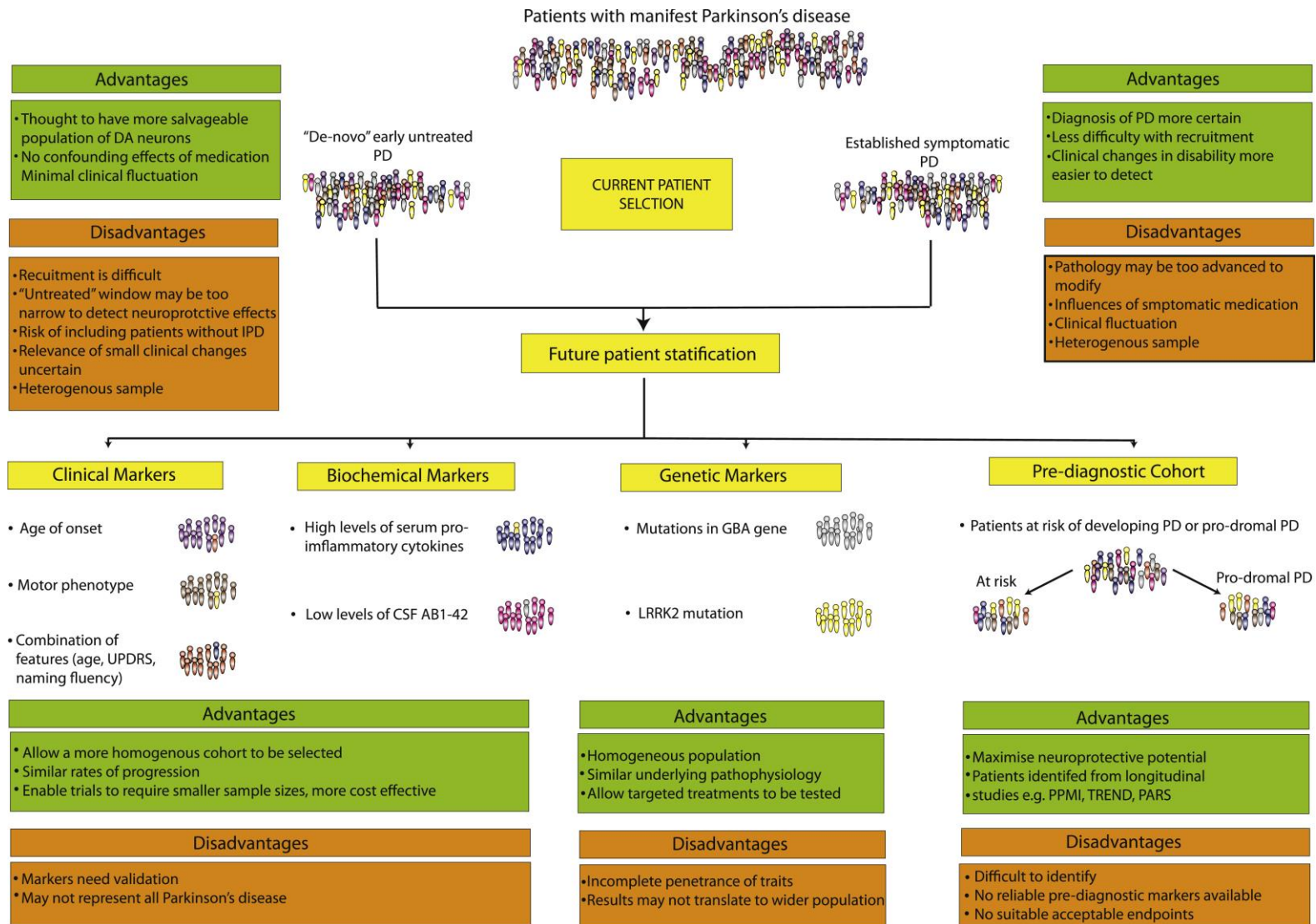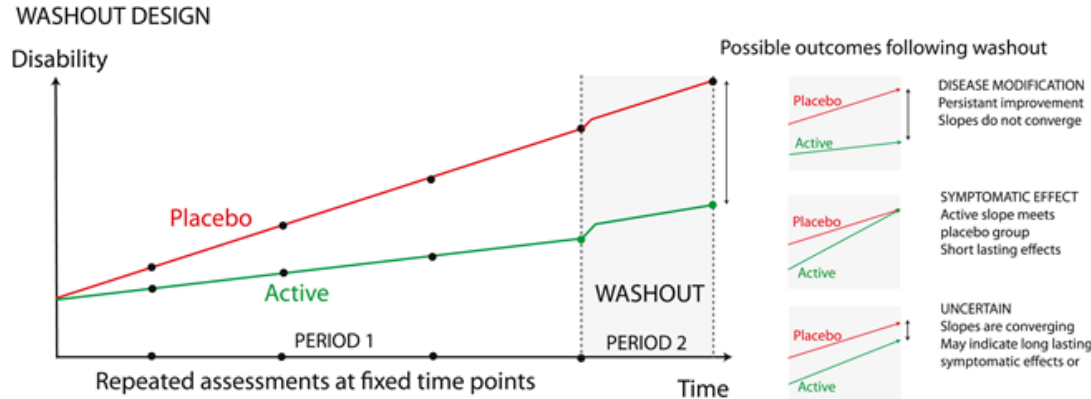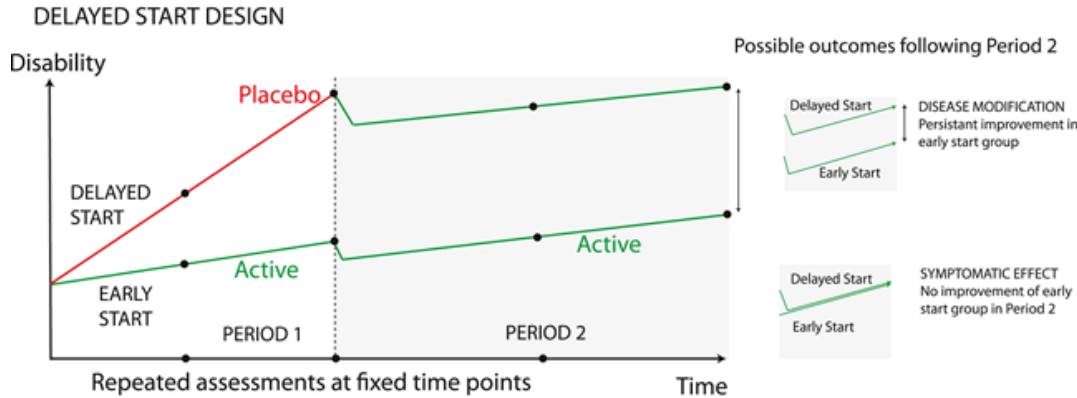

Patients with manifest Parkinson's disease

Advantages
- Thought to have more salvageable population of DA neurons
- No confounding effects of medication Minimal clinical fluctuation

"De-novo" early untreated PD

CURRENT PATIENT SELCTION

Established symptomatic PD

Advantages
- Diagnosis of PD more certain
- Less difficulty with recruitment
- Clinical changes in disability more easier to detect

Disadvantages
- Recuitment is difficult
- "Untreated" window may be too narrow to detect neuroprotctive effects
- Risk of including patients without IPD
- Relevance of small clinical changes uncertain
- Heterogenous sample

Disadvantages
- Pathology may be too advanced to modify
- Influences of smptomatic medication
- Clinical fluctuation
- Heterogenous sample

Future patient statification

Clinical Markers
- Age of onset
- Motor phenotype
- Combination of features (age, UPDRS, naming fluency)

Biochemical Markers
- High levels of serum pro-imflammatory cytokines
- Low levels of CSF AB1-42

Genetic Markers
- Mutations in GBA gene
- LRRK2 mutation

Pre-diagnostic Cohort
- Patients at risk of developing PD or pro-dromal PD

At risk   Pro-dromal PD

Advantages
- Allow a more homogenous cohort to be selected
- Similar rates of progression
- Enable trials to require smaller sample sizes, more cost effective

Advantages
- Homogeneous population
- Similar underlying pathophysiology
- Allow targeted treatments to be tested

Advantages
- Maximise neuroprotective potential
- Patients identifed from longitudinal
- studies e.g. PPMI, TREND, PARS

Disadvantages
- Markers need validation
- May not represent all Parkinson's disease

Disadvantages
- Incomplete penetrance of traits
- Results may not translate to wider population

Disadvantages
- Difficult to identify
- No reliable pre-diagnostic markers available
- No suitable acceptable endpoints

Figure 2 - Common trial designs used in Phase 3 disease modification trials.

References

[1]   C.W. Olanow, K. Kieburtz, A.H. V Schapira, Why have we failed to achieve neuroprotection in Parkinson's disease?, Ann. Neurol. 64 Suppl 2 (2008) S101–10. doi:10.1002/ana.21461.

[2]   K. Kieburtz, C.W. Olanow, Advances in clinical trials for movement disorders., Mov. Disord. 30 (2015) 1580–7. doi:10.1002/mds.26371.

[3]   O. Rascol, "Disease-modification" trials in Parkinson disease: target populations, endpoints and study design., Neurology. 72 (2009) S51–8. doi:10.1212/WNL.0b013e318199049e.

[4]   C.J. Van der Schyf, Rational drug discovery design approaches for treating Parkinson's disease., Expert Opin. Drug Discov. 10 (2015) 713–41. doi:10.1517/17460441.2015.1041495.

[5]   G. Jin, S.T.C. Wong, Toward better drug repositioning: prioritizing and integrating existing methods into efficient pipelines., Drug Discov. Today. 19 (2014) 637–44. doi:10.1016/j.drudis.2013.11.005.

[6]   A.H. Rajput, A. Voll, M.L. Rajput, C.A. Robinson, A. Rajput, Course in Parkinson disease subtypes: A 39-year clinicopathologic study., Neurology. 73 (2009) 206–12. doi:10.1212/WNL.0b013e3181ae7af1.

[7]   A.H. Rajput, H.H. Sitte, A. Rajput, M.E. Fenton, C. Pifl, O. Hornykiewicz, Globus pallidus dopamine and Parkinson motor subtypes: clinical and brain biochemical correlation., Neurology. 70 (2008) 1403–10. doi:10.1212/01.wnl.0000285082.18969.3a.

[8]   G. Reinoso, J.C. Allen, W.L. Au, S.H. Seah, K.Y. Tay, L.C.S. Tan, Clinical evolution of Parkinson's disease and prognostic factors affecting motor progression: 9-year follow-up study, Eur. J. Neurol. 22 (2015) 457–463. doi:10.1111/ene.12476.

[9]   G. Pagano, N. Ferrara, D.J. Brooks, Age at onset and Parkinson disease phenotype, Neurology. (2016). doi:10.1212/WNL.0000000000002461.

[10]   T. Foltynie, C. Brayne, R.A. Barker, The heterogeneity of idiopathic Parkinson's disease., J. Neurol. 249 (2002)

138–45. http://www.ncbi.nlm.nih.gov/pubmed/11985378 (accessed July 25, 2014).

[11]   S. Fereshtehnejad, S. Romenets, J. Anang, et al., New Clinical Subtypes of Parkinson Disease and Their

Longitudinal Progression, JAMA Neurol. 72 (2015) E1–11. doi:10.1001/jamaneurol.2015.0703.

[12]   D.C. Velseboer, R.M.A. de Bie, L. Wieske, J.R. Evans, S.L. Mason, T. Foltynie, et al., Development and external

validation of a prognostic model in newly diagnosed Parkinson disease., Neurology. 86 (2016) 986–93.

doi:10.1212/WNL.0000000000002437.

[13]   T. Simuni, C. Caspell-Garcia, C. Coffey, S. Lasch, C. Tanner, K. Marek, How stable are Parkinson's disease

subtypes in de novo patients: Analysis of the PPMI cohort?, Parkinsonism Relat. Disord. 28 (2016) 62–67.

doi:10.1016/j.parkreldis.2016.04.027.

[14]   J.-H. Kang, B. Mollenhauer, C.S. Coffey, J.B. Toledo, D. Weintraub, D.R. Galasko, et al., CSF biomarkers

associated with disease heterogeneity in early Parkinson's disease: the Parkinson's Progression Markers

Initiative study - Springer, Acta Neuropathol. (2016). http://link.springer.com/article/10.1007/s00401-016-

1552-2/fulltext.html#CR30 (accessed April 27, 2016).

[15]   C.H. Williams-Gray, R. Wijeyekoon, A.J. Yarnall, R.A. Lawson, D.P. Breen, J.R. Evans, et al., Serum immune

markers and disease progression in an incident Parkinson's disease cohort (ICICLE-PD), Mov. Disord. 00 (2016)

n/a–n/a. doi:10.1002/mds.26563.

[16]   O. Corti, M. Fournier, A. Brice, Neurodegeneration in Parkinson's disease: genetics enlightens physiopathology.,

J. Neural Transm. Suppl. (2009) 215–21. http://www.ncbi.nlm.nih.gov/pubmed/20411780 (accessed April 25,

2016).

[17]   L. Pihlstrøm, K.R. Morset, E. Grimstad, V. Vitelli, A cumulative genetic risk score predicts motor progression in

Parkinson ' s disease, Mov. Disord. 00 (2016) 1–8. doi:10.1002/mds.26505.

[18]   O. Cooper, H. Seo, S. Andrabi, C. Guardia-Laguarta, J. Graziotto, M. Sundberg, et al., Pharmacological rescue of

mitochondrial deficits in iPSC-derived neural cells from patients with familial Parkinson's disease., Sci. Transl.

Med. 4 (2012) 141ra90. doi:10.1126/scitranslmed.3003985.

[19]   M. Rasool, A. Malik, M.I. Naseer, A. Manan, S. Ansari, I. Begum, et al., The role of epigenetics in personalized

medicine: challenges and opportunities., BMC Med. Genomics. 8 Suppl 1 (2015) S5. doi:10.1186/1755-8794-8-S1-S5.

[20]   A. McNeill, J. Magalhaes, C. Shen, K.-Y. Chau, D. Hughes, A. Mehta, et al., Ambroxol improves lysosomal biochemistry in glucocerebrosidase mutation-linked Parkinson disease cells., Brain. 137 (2014) 1481–95. doi:10.1093/brain/awu020.

[21]   J.H. Kordower, C.W. Olanow, H.B. Dodiya, Y. Chu, T.G. Beach, C.H. Adler, et al., Disease duration and the integrity of the nigrostriatal system in Parkinson's disease., Brain. 136 (2013) 2419–31. doi:10.1093/brain/awt192.

[22]   E. Tolosa, C. Gaig, J. Santamaría, Y. Compta, Diagnosis and the premotor phase of Parkinson disease., Neurology. 72 (2009) S12–20. doi:10.1212/WNL.0b013e318198db11.

[23]   R. Hilker, K. Schweitzer, S. Coburger, M. Ghaemi, S. Weisenbach, A.H. Jacobs, et al., Nonlinear progression of Parkinson disease as determined by serial positron emission tomographic imaging of striatal fluorodopa F 18 activity., Arch. Neurol. 62 (2005) 378–82. doi:10.1001/archneur.62.3.378.

[24]   A.J. Noyce, J.P. Bestwick, L. Silveira-Moriyama, C.H. Hawkes, G. Giovannoni, A.J. Lees, et al., Meta-analysis of early nonmotor features and risk factors for Parkinson disease., Ann. Neurol. 72 (2012) 893–901. doi:10.1002/ana.23687.

[25]   A.J. Noyce, A.J. Lees, A.-E. Schrag, The prediagnostic phase of Parkinson's disease, J. Neurol. Neurosurg. Psychiatry. (2016) jnnp–2015–311890. doi:10.1136/jnnp-2015-311890.

[26]   D. Salat, A.J. Noyce, A. Schrag, E. Tolosa, Challenges of modifying disease progression in prediagnostic Parkinson's disease, Lancet Neurol. 4422 (2016) 1–12. doi:10.1016/S1474-4422(16)00060-0.

[27]   M.R. Keezer, C. Wolfson, R.B. Postuma, Age, Gender, Comorbidity, and the MDS-UPDRS: Results from a Population-Based Study., Neuroepidemiology. 46 (2016) 222–227. doi:10.1159/000444021.

[28]   D. Berg, S. Behnke, K. Seppi, J. Godau, S. Lerche, P. Mahlknecht, et al., Enlarged hyperechogenic substantia nigra as a risk marker for Parkinson's disease., Mov. Disord. 28 (2013) 216–9. doi:10.1002/mds.25192.

[29]   B. Dehay, Targeting α-synuclein: Therapeutic options, Mov. Disord. (2016).

http://www.ncbi.nlm.nih.gov/pubmed/26926119 (accessed April 26, 2016).

[30] M. Hashimoto, L.J. Hsu, E. Rockenstein, T. Takenouchi, M. Mallory, E. Masliah, alpha-Synuclein protects against oxidative stress via inactivation of the c-Jun N-terminal kinase stress-signaling pathway in neuronal cells., J. Biol. Chem. 277 (2002) 11465–72. doi:10.1074/jbc.M111428200.

[31] C.W. Olanow, R.A. Hauser, J. Jankovic, W. Langston, A. Lang, W. Poewe, et al., A randomized, double-blind, placebo-controlled, delayed start study to assess rasagiline as a disease modifying therapy in Parkinson's disease (the ADAGIO study): rationale, design, and baseline characteristics., Mov. Disord. 23 (2008) 2194–201. doi:10.1002/mds.22218.

[32] S. Fahn, D. Oakes, I. Shoulson, K. Kieburtz, A. Rudolph, A. Lang, et al., Levodopa and the progression of Parkinson's disease., N. Engl. J. Med. 351 (2004) 2498–508. doi:10.1056/NEJMoa033447.

[33] C.W. Olanow, R.A. Hauser, L. Gauger, T. Malapira, W. Koller, J. Hubble, et al., The effect of deprenyl and levodopa on the progression of Parkinson's disease., Ann. Neurol. 38 (1995) 771–7. doi:10.1002/ana.410380512.

[34] R.A. Hauser, N.H.G. Holford, Quantitative description of loss of clinical benefit following withdrawal of levodopa-carbidopa and bromocriptine in early Parkinson's disease., Mov. Disord. 17 (2002) 961–8. doi:10.1002/mds.10226.

[35] T.A. Mestre, P. Shah, C. Marras, G. Tomlinson, A.E. Lang, Another face of placebo: the lessebo effect in Parkinson disease: meta-analyses., Neurology. 82 (2014) 1402–9. doi:10.1212/WNL.0000000000000340.

[36] H. Liu-Seifert, S.W. Andersen, I. Lipkovich, K.C. Holdridge, E. Siemers, A novel approach to delayed-start analyses for demonstrating disease-modifying effects in Alzheimer's disease., PLoS One. 10 (2015) e0119632. doi:10.1371/journal.pone.0119632.

[37] Parkinson Study Group, A controlled trial of rasagiline in early Parkinson disease: the TEMPO Study., Arch. Neurol. 59 (2002) 1937–43. http://www.ncbi.nlm.nih.gov/pubmed/12470183 (accessed February 22, 2016).

[38] A.H. V Schapira, M.P. McDermott, P. Barone, C.L. Comella, S. Albrecht, H.H. Hsu, et al., Pramipexole in patients with early Parkinson's disease (PROUD): a randomised delayed-start trial., Lancet. Neurol. 12 (2013) 747–55.

doi:10.1016/S1474-4422(13)70117-0.

[39]   D. Devos, C. Moreau, J.C. Devedjian, J. Kluza, M. Petrault, C. Laloux, et al., Targeting chelatable iron as a

therapeutic modality in Parkinson's disease., Antioxid. Redox Signal. 21 (2014) 195–210.

doi:10.1089/ars.2013.5593.

[40]   C.V.M. Verschuur, S.R. Suwijn, B. Post, M. Dijkgraaf, B.R. Bloem, J.J. van Hilten, et al., Protocol of a randomised

delayed-start double-blind placebo-controlled multi-centre trial for Levodopa in EArly Parkinson's disease: the

LEAP-study., BMC Neurol. 15 (2015) 236. doi:10.1186/s12883-015-0491-1.

[41]   L.M. Shulman, A.L. Gruber-Baldini, K.E. Anderson, P.S. Fishman, S.G. Reich, W.J. Weiner, The clinically important

difference on the unified Parkinson's disease rating scale., Arch. Neurol. 67 (2010) 64–70.

doi:10.1001/archneurol.2009.295.

[42]   A. Schrag, C. Sampaio, N. Counsell, W. Poewe, Minimal clinically important change on the unified Parkinson's

disease rating scale., Mov. Disord. 21 (2006) 1200–7. doi:10.1002/mds.20914.

[43]   O. Rascol, Defining a minimal clinically relevant difference for the unified Parkinson's rating scale: an important

but still unmet need., Mov. Disord. 21 (2006) 1059–61. doi:10.1002/mds.20913.

[44]   S.R. Schwid, G.R. Cutter, Editorial CME Futility studies Spending a little to save a lot, Neurology. 66 (2006) 626–

627.

[45]   B. Levin, The futility study--Progress over the last decade., Contemp. Clin. Trials. 45 (2015) 69–75.

doi:10.1016/j.cct.2015.06.013.

[46]   NINDS Exploratory Trials in Parkinson Disease (NET-PD) FS-ZONE Investigators., Pioglitazone in early Parkinson's

disease: a phase 2, multicentre, double-blind, randomised trial., Lancet. Neurol. 14 (2015) 795–803.

doi:10.1016/S1474-4422(15)00144-1.

[47]   S.A. Parashos, S. Luo, K.M. Biglan, I. Bodis-Wollner, B. He, G.S. Liang, et al., Measuring disease progression in

early Parkinson disease: the National Institutes of Health Exploratory Trials in Parkinson Disease (NET-PD)

experience., JAMA Neurol. 71 (2014) 710–6. doi:10.1001/jamaneurol.2014.391.

[48]   Mixed lineage kinase inhibitor CEP-1347 fails to delay disability in early Parkinson disease., Neurology. 69

(2007) 1480–90. doi:10.1212/01.wnl.0000277648.63931.c0.

[49] A.E. Lang, E. Melamed, W. Poewe, O. Rascol, Trial designs used to study neuroprotective therapy in Parkinson's disease., Mov. Disord. 28 (2013) 86–95. doi:10.1002/mds.24997.

[50] I. Aviles-Olmos, J. Dickson, Z. Kefalopoulou, A. Djamshidian, P. Ell, T. Soderlund, et al., Exenatide and the treatment of patients with Parkinson ' s disease, 123 (2013). doi:10.1172/JCI68295.2730.

[51] I. Aviles-Olmos, J. Dickson, Z. Kefalopoulou, A. Djamshidian, J. Kahan, P.E. Fmedsci, et al., Motor and Cognitive Advantages Persist 12 Months After Exenatide Exposure in Parkinson's Disease., J. Parkinsons. Dis. (2014). doi:10.3233/JPD-140364.

[52] R.A. Barker, M. Stacy, P. Brundin, A new approach to disease-modifying drug trials in Parkinson's disease., J. Clin. Invest. 123 (2013) 2364–5. doi:10.1172/JCI69690.

[53] R. de la Fuente-Fernández, A.J. Stoessl, The placebo effect in Parkinson's disease, Trends Neurosci. 25 (2002) 302–306. doi:10.1016/S0166-2236(02)02181-1.

[54] S.C. Lidstone, Great expectations: the placebo effect in Parkinson's disease., Handb. Exp. Pharmacol. 225 (2014) 139–47. doi:10.1007/978-3-662-44519-8_8.

[55] A.E. Lang, S. Gill, N.K. Patel, A. Lozano, J.G. Nutt, R. Penn, et al., Randomized controlled trial of intraputamenal glial cell line-derived neurotrophic factor infusion in Parkinson disease., Ann. Neurol. 59 (2006) 459–66. doi:10.1002/ana.20737.

[56] M. Hutchinson, S. Gurney, R. Newson, GDNF in Parkinson disease: an object lesson in the tyranny of type II., J. Neurosci. Methods. 163 (2007) 190–2. doi:10.1016/j.jneumeth.2006.06.015.

[57] B. Mcfadden, L. Heitzman-powell, Novel Methods and Technologies for 21st-Century Clinical Trials: a review, JAMA Neurol. 8 (2015) 1699–1712. doi:10.1016/j.rasd.2014.08.015.Social.

[58] B.A. Ploeger, N.H.G. Holford, Washout and delayed start designs for identifying disease modifying effects in slowly progressive diseases using disease progression analysis., Pharm. Stat. 8 (2009) 225–38. doi:10.1002/pst.355.

[59] M.A. Björnsson, L.E. Friberg, U.S.H. Simonsson, Performance of nonlinear mixed effects models in the presence

of informative dropout., AAPS J. 17 (2015) 245–55. doi:10.1208/s12248-014-9700-x.

[60]   O. Siddiqui, H.M.J. Hung, R. O'Neill, MMRM vs. LOCF: a comprehensive comparison based on simulation study and 25 NDA datasets., J. Biopharm. Stat. 19 (2009) 227–46. doi:10.1080/10543400802609797.

[61]   C.W. Olanow, A.H. V Schapira, P.A. LeWitt, K. Kieburtz, D. Sauer, G. Olivieri, et al., TCH346 as a neuroprotective drug in Parkinson's disease: a double-blind, randomised, controlled trial., Lancet Neurol. 5 (2006) 1013–20. doi:10.1016/S1474-4422(06)70602-0.

[62]   S.A. Parashos, C.J. Swearingen, K.M. Biglan, I. Bodis-Wollner, G.S. Liang, G.W. Ross, et al., Determinants of the timing of symptomatic treatment in early Parkinson disease: The National Institutes of Health Exploratory Trials in Parkinson Disease (NET-PD) Experience., Arch. Neurol. 66 (2009) 1099–104. doi:10.1001/archneurol.2009.159.

[63]   C.D. Ward, Does selegiline delay progression of Parkinson's disease? A critical re-evaluation of the DATATOP study., J. Neurol. Neurosurg. Psychiatry. 57 (1994) 217–20. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1072455&tool=pmcentrez&rendertype=abstract (accessed August 11, 2014).

[64]   P.A. Kempster, S.S. O'Sullivan, J.L. Holton, T. Revesz, A.J. Lees, Relationships between age and late progression of Parkinson's disease: a clinico-pathological study., Brain. 133 (2010) 1755–62. doi:10.1093/brain/awq059.

[65]   C.H. Williams-Gray, S.L. Mason, J.R. Evans, T. Foltynie, C. Brayne, T.W. Robbins, et al., The CamPaIGN study of Parkinson's disease: 10-year outlook in an incident population-based cohort., J. Neurol. Neurosurg. Psychiatry. 84 (2013) 1258–64. doi:10.1136/jnnp-2013-305277.

[66]   J. Evans, Disease-modifying therapy trials in PD: what are the issues?, ACNR. (2014) 14(3): 6–8. http://www.acnr.co.uk/2014/08/disease-modifying-therapy-trials-in-pd-what-are-the-issues/ (accessed August 29, 2014).

[67]   D. McGhee, A. Parker, S. Fielding, J. Zajicek, C. Counsell, Using "dead or dependent" as an outcome measure in clinical trials in Parkinson's disease., J. Neurol. Neurosurg. Psychiatry. 86 (2015) 180–5. doi:10.1136/jnnp-2014-307703.

[68]   C.G. Goetz, B.C. Tilley, S.R. Shaftman, G.T. Stebbins, S. Fahn, P. Martinez-Martin, et al., Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results., Mov. Disord. 23 (2008) 2129–70. doi:10.1002/mds.22340.

[69]   C.G. Goetz, G.T. Stebbins, B.C. Tilley, Calibration of unified Parkinson's disease rating scale scores to Movement Disorder Society-unified Parkinson's disease rating scale scores., Mov. Disord. 27 (2012) 1239–42. doi:10.1002/mds.25122.

[70]   J. Williams, Z. Mari, G. Pontone, S. Bassett, Natural History of UPDRS Motor Scores in an Observational Parkinson's Disease Cohort (S22.004), Neurology. 78 (2012) S22.004–. http://www.neurology.org/content/78/1_Supplement/S22.004.short (accessed April 15, 2016).

[71]   S.S. Papapetropoulos, Patient diaries as a clinical endpoint in Parkinson's disease clinical trials., CNS Neurosci. Ther. 18 (2012) 380–7. doi:10.1111/j.1755-5949.2011.00253.x.

[72]   H. Utsumi, H. Terashi, Y. Ishimura, T. Takazawa, Y. Okuma, M. Yoneyama, et al., How far do the complaints of patients with Parkinson's disease reflect motor fluctuation? Quantitative analysis using a portable gait rhythmogram., ISRN Neurol. 2012 (2012) 372030. doi:10.5402/2012/372030.

[73]   M.B. Davidson, D.J.M. McGhee, C.E. Counsell, Comparison of patient rated treatment response with measured improvement in Parkinson's disease., J. Neurol. Neurosurg. Psychiatry. 83 (2012) 1001–5. doi:10.1136/jnnp-2012-302741.

[74]   D. McGhee, A. Parker, S. Fielding, C. Counsell, WHICH CLINICAL MEASURES ARE MOST APPROPRIATE FOR MEASURING DISEASE PROGRESSION IN PARKINSON'S DISEASE?, J. Neurol. Neurosurg. Psychiatry. 84 (2013) e2–e2. doi:10.1136/jnnp-2013-306573.163.

[75]   D.J. Brooks, N. Pavese, Imaging biomarkers in Parkinson's disease., Prog. Neurobiol. 95 (2011) 614–28. doi:10.1016/j.pneurobio.2011.08.009.

[76]   G.D. Femminella, P. Edison, Evaluation of neuroprotective effect of glucagon-like peptide 1 analogs using neuroimaging., Alzheimers. Dement. 10 (2014) S55–S61. doi:10.1016/j.jalz.2013.12.012.

[77]   A.L. Whone, R.L. Watts, A.J. Stoessl, M. Davis, S. Reske, C. Nahmias, et al., Slower progression of Parkinson's

disease with ropinirole versus levodopa: The REAL-PET study., Ann. Neurol. 54 (2003) 93–101. doi:10.1002/ana.10609.

[78]   B. Ravina, D. Eidelberg, J.E. Ahlskog, R.L. Albin, D.J. Brooks, M. Carbon, et al., The role of radiotracer imaging in Parkinson disease., Neurology. 64 (2005) 208–15. doi:10.1212/01.WNL.0000149403.14458.7F.

[79]   C.R. Freed, P.E. Greene, R.E. Breeze, W.-Y. Tsai, W. DuMouchel, R. Kao, et al., Transplantation of Embryonic Dopamine Neurons for Severe Parkinson's Disease, N. Engl. J. Med. 344 (2001) 710–719. doi:10.1056/NEJM200103083441002.

[80]   A.D. Andersen, M. Binzer, E. Stenager, J.B. Gramsbergen, Cerebrospinal fluid biomarkers for Parkinson's disease - a systematic review., Acta Neurol. Scand. (2016). doi:10.1111/ane.12590.

[81]   D.B. Miller, J.P. O'Callaghan, Biomarkers of Parkinson's Disease (Pd): Present and Future, Metabolism. 64 (2014) S40–6. doi:10.1016/j.metabol.2014.10.030.

[82]   D.P. Bagchi, L. Yu, J.S. Perlmutter, J. Xu, R.H. Mach, Z. Tu, et al., Binding of the radioligand SIL23 to α-synuclein fibrils in Parkinson disease brain tissue establishes feasibility and screening approaches for developing a Parkinson disease imaging agent., PLoS One. 8 (2013) e55031. doi:10.1371/journal.pone.0055031.

[83]   E. Mak, L. Su, G.B. Williams, M.J. Firbank, R.A. Lawson, A.J. Yarnall, et al., Baseline and longitudinal grey matter changes in newly diagnosed Parkinson's disease: ICICLE-PD study., Brain. 138 (2015) 2974–86. doi:10.1093/brain/awv211.

[84]   B. Post, M.P. Merkus, R.M.A. de Bie, R.J. de Haan, J.D. Speelman, Unified Parkinson's disease rating scale motor examination: are ratings of nurses, residents in neurology, and movement disorders specialists interchangeable?, Mov. Disord. 20 (2005) 1577–84. doi:10.1002/mds.20640.

[85]   L. Newington, A. Metcalfe, Factors influencing recruitment to research: qualitative study of the experiences and perceptions of research teams., BMC Med. Res. Methodol. 14 (2014) 10. doi:10.1186/1471-2288-14-10.

[86]   A. Kott, J. Swartz, Does Rater Experience Influence UPDRS Inter-rater Reliability?, in: 16th Int. Congr. Park. Dis. Mov. Disord., 2012. https://www.bracketglobal.com/wp-content/uploads/DoesRaterExperienceInfluenceUPDRSInterRaterReliabilityMDS2012.pdf (accessed April 20,

2016).

[87]   C. Ossig, Wearable sensor-based objective assessment of motor symptoms in Parkinson's disease - Springer, J. Neural Transm. (2016). http://link.springer.com/article/10.1007/s00702-015-1439-8/fulltext.html (accessed April 20, 2016).

[88]   S. Arora, V. Venkataraman, A. Zhan, S. Donohue, K.M. Biglan, E.R. Dorsey, et al., Detecting and monitoring the symptoms of Parkinson's disease using smartphones: A pilot study., Parkinsonism Relat. Disord. 21 (2015) 650–3. doi:10.1016/j.parkreldis.2015.02.026.

[89]   A. Abdolahi, N. Scoglio, A. Killoran, E.R. Dorsey, K.M. Biglan, Potential reliability and validity of a modified version of the Unified Parkinson's Disease Rating Scale that could be administered remotely., Parkinsonism Relat. Disord. 19 (2013) 218–21. doi:10.1016/j.parkreldis.2012.10.008.

[90]   Q.W. Oung, H. Muthusamy, H.L. Lee, S.N. Basah, S. Yaacob, M. Sarillee, et al., Technologies for Assessment of Motor Disorders in Parkinson's Disease: A Review., Sensors (Basel). 15 (2015) 21710–45. doi:10.3390/s150921710.

[91]   A.T. Tzallas, M.G. Tsipouras, G. Rigas, D.G. Tsalikakis, E.C. Karvounis, M. Chondrogiorgi, et al., PERFORM: a system for monitoring, assessment and management of patients with Parkinson's disease., Sensors (Basel). 14 (2014) 21329–57. doi:10.3390/s141121329.

[92]   T.O. Mera, D.A. Heldman, A.J. Espay, M. Payne, J.P. Giuffrida, Feasibility of home-based automated Parkinson's disease motor assessment., J. Neurosci. Methods. 203 (2012) 152–6. doi:10.1016/j.jneumeth.2011.09.019.

[93]   E.R. Dorsey, V. Venkataraman, M.J. Grana, M.T. Bull, B.P. George, C.M. Boyd, et al., Randomized controlled clinical trial of "virtual house calls" for Parkinson disease., JAMA Neurol. 70 (2013) 565–70. doi:10.1001/jamaneurol.2013.123.

[94]   D.A. Heldman, A.J. Espay, P.A. LeWitt, J.P. Giuffrida, Clinician versus machine: reliability and responsiveness of motor endpoints in Parkinson's disease., Parkinsonism Relat. Disord. 20 (2014) 590–5. doi:10.1016/j.parkreldis.2014.02.022.