Contemporary Clinical Trials Communications 15 (2019) 100397

Contents lists available at ScienceDirect

Contemporary Clinical Trials Communications

journal homepage: www.elsevier.com/locate/conctc

Protocol: Using N-of-1 tests to identify responders to melatonin for sleep disturbance in Parkinson's disease

J. Nikles[a,*], J.D. O'Sullivan[b], G.K. Mitchell[c], S.S. Smith[d], J.M. McGree[f], H. Senior[e], N. Dissanyaka[g], A. Ritchie[h]

[a] UQ Centre for Clinical Research, The University of Queensland, Brisbane, Australia
[b] Royal Brisbane & Women's Hospital (RBWH), Neurology Department, School of Medicine, St Andrew's, Wesley and RBWH Hospitals, The University of Queensland, Brisbane, Australia
[c] General Practice and Palliative Care, School of Medicine, The University of Queensland, Brisbane, Australia
[d] Institute for Social Science Research (ISSR), The University of Queensland, Brisbane, Australia
[e] Human Biology and Health, College of Health, Massey University, Auckland, New Zealand
[f] School of Mathematical Sciences, Queensland University of Technology, Brisbane, Australia
[g] The University of Queensland, Brisbane, Australia
[h] St Andrew's War Memorial Hospital, Royal Brisbane and Women's Hospital, Brisbane, Australia

ARTICLE INFO

Keywords:
N-of-1 trial
Parkinson's disease
Melatonin
Sleep
Insomnia

ABSTRACT

Background: 40% of Parkinson's Disease (PD) sufferers experience insomnia, impacting health and quality of life for patients and family members, especially carers. There is little evidence that current treatments are effective.
Objectives: To determine the effectiveness of melatonin in reducing insomnia in 44 individuals with PD using N-of-1 trials. To aggregate group data to arrive at population estimates of effectiveness (measured by improvements in PDSS-2) and safety (measured by adverse events) of melatonin in improving insomnia in PD. To assess the feasibility of offering N-of-1 trials for insomnia in PD.
Methodology: Participants will receive either immediate-release melatonin or placebo in random order in 3 paired two-week treatment periods (12 weeks total). Based on their response in a two-week run-in period on 3 mg daily, they will trial either 3 mg or 6 mg. Patients will keep daily sleep diaries and wear a MotionWatch throughout. After the trial patients will discuss their individual report with their doctor, which provides direct feedback about effectiveness and safety of melatonin for them.
Statistical methods: We will analyse N-of-1 tests 1) individually: effects of melatonin on PDSS-2 and safety will be reported; and 2) aggregated across individual N-of-1 studies, combined using a Bayesian multilevel random effects model, which will account for repeated measures on individuals over time, and will return posterior estimates of overall treatment effect, and effect in each individual.
Clinical Trial Registration number: ACTRN12617001103358.

1. Background

Australia's second most common neurological disease, Parkinson's Disease (PD), has grown by 17% over the last six years, costing Australia $7.6 billion in 2012 [1]. Two-thirds of patients with PD experience one or more sleep-related non-motor symptoms including insomnia, REM-sleep behaviour disorder (RBD), excessive daytime sleepiness, restless legs syndrome, and obstructive sleep apnoea [2]. Insomnia is the most common of these complaints [3] affecting 37% of PD patients [2]. Brainstem neurodegenerative processes, disturbances of circadian rhythm, the effect of symptoms of PD on sleep and concomitant sleep disorders contribute to sleep disturbances [4].

Although there is a well established correlation between anxiety and sleep disturbances in PD [5], there is little evidence for therapeutic interventions for sleep problems in PD. Cognitive Behaviour Therapy is recommended for insomnia but there has only been one unblinded study in PD specifically, where it was combined with light therapy and it was not possible to differentiate beneficial effects of the interventions and some measures, including quality of life, actually deteriorated [6]. The Movement Disorder Society review [7] concluded there is insufficient data to recommend any specific long-term treatment of sleep-related problems in PD.

* Corresponding author.
E-mail address: uqjnikle@uq.edu.au (J. Nikles).

https://doi.org/10.1016/j.conctc.2019.100397
Received 10 January 2019; Received in revised form 5 June 2019; Accepted 13 June 2019
Available online 08 July 2019

Individuals with neurological disorders often have circadian rhythm disturbances [8]. PD has many features exhibiting diurnal fluctuations, suggestive of circadian involvement [9]. Melatonin may improve sleep in PD as a chronobiotic or as a hypnotic.

In other populations melatonin can increase sleep efficiency, decrease night-time activity, shorten sleep onset latency and improve insomnia [10]. Recent systematic reviews [3,11] concluded that melatonin improves insomnia but not objective sleep outcomes in PD patients and the available studies were small (n = 40 and 18) and short-term (2 and 4 weeks).

Although these reviews suggest that melatonin is relatively safe and that use in people with PD is based on sound theoretical considerations, a longer term well-designed trial of sufficient sample size is needed [12] to determine the benefit of melatonin for insomnia in PD patients. Melatonin is readily available despite this lack of evidence. The response rate to melatonin in people with PD and insomnia is also unknown. Not all older adults respond to melatonin [13], so identifying responders would be efficient and novel.

N-of-1 designs, clinical trials involving single patients serving as their own control, determine response and benefit for each individual.

N-of-1 trials are ethical, multiple crossover, randomised controlled trials of interventions versus a placebo or another intervention[1], which test the effect of an intervention in an individual patient. In N-of-1 trials, a patient crosses back and forth between two treatments in random order, several times, enabling clinicians to identify the more effective intervention for that individual, with greater precision than can generally be achieved by using the informal trial-and-error paradigm commonly used in clinical practice to guide treatment plans.[2,3] The effect of a treatment of interest and a comparator is measured within each of the multiple cycles (i.e. a pair of treatments in each cycle; Fig. 1). N-of-1 trials have emerged as a high-integrity, evidence-based, pragmatic clinical decision methodology for patient-centered comparative effectiveness research, particularly if response is variable, the treatment is expensive or has significant side effects, the condition is rare, or recruitment or retention are difficult Aggregated N-of-1 trials can reduce the sample size required, compared to an RCT, to produce group estimates of effect, because of greater statistical power [14], with implications for study design in low prevalence diseases such as PD (0.3% of general population [15].

N-of-1 trials provide the highest level of evidence possible, identifying the most effective intervention, and helping to identify which treatments may cause more harm than good for individual patients. Thus, therapy is optimised by ensuring that the 'right treatment is given to the right patient', and is continued only in those who respond, thereby reducing harm, improving health outcomes, informing prescribing decisions and improving cost effectiveness.

1.1. Aims

The primary aim of this study is to determine the effectiveness of melatonin in reducing insomnia in people with PD, based on change in PDSS-2 scores.

Fig. 1. Example order of treatment periods.

1.1.1. Secondary aims are

1. To use N-of-1 trials (tests) to improve the precision of clinical decision making in prescribing melatonin for patients with PD and insomnia, by identifying if they respond to melatonin.
2. To aggregate group data from a series of N-of-1 trials (tests) to arrive at population estimates of the effectiveness (measured by improvements in sleep onset latency (SOL), sleep efficiency and sleep-related impairment) and safety (measured by adverse events) of melatonin in improving insomnia in PD.
3. To assess the feasibility of offering N-of-1 trials (tests) for insomnia to suitable people with PD.

1.2. Primary study endpoint

Insomnia (sleep problems) measured by the Parkinson's Disease Sleep Scale 2nd version (PDSS-2; [16]. We will use a threshold of −3.44 points change in the PDSS-2 for detecting a clinically significant improvement [17].

1.3. Study design

Although RCTs are the gold standard for evaluating treatment efficacy, an alternative is an N-of-1 trial. N-of-1 trials, or tests, use a multicycle, double-blind, randomised controlled trial (RCT) design where each participant is assured of receiving both the study medication and placebo, and thus learns whether the treatment works specifically for them. N-of-1 tests have been previously used in PD to test caffeine for daytime somnolence [18] and naftazone for antiparkinsonian and antidyskinetic effects [19]. Melatonin is an optimal treatment for testing within an N-of-1 trial as: (i) it has a short half-life (36–41 min); (ii) there is no residual impact on the target symptom after excretion; and (iii) it is being used to treat an important and recurrent symptom that has a negative impact on quality of life (QoL) in PD [2].

Each participant will undergo 3 pairs of treatment/placebo periods (each period is 2 weeks; each N-of-1 test will last 12 weeks). Treatments will be randomised separately within each pair of periods (e.g. ABBAAB) by the study statistician. Patients, clinicians, research staff and outcome assessors will be blinded. Data from day 1 will be discarded to provide a washout period. If a carry-over effect is found subsequent to day 1, appropriate adjustments will be made. Results from the individual N-of-1 tests will be pooled for meta-analysis. Fig. 1 shows a schematic representation of the design.

1.4. Eligibility criteria

1.4.1. Inclusion criteria

1. Adult patients 30 years or more with a diagnosis of idiopathic PD according to the UK Brain Bank [20].
2. Patient has claimed chronic sleep difficulty which is impacting his/her life.
3. Score > 5 on the Pittsburgh Sleep Quality Index (PSQI) [21].
4. If on sedatives or hypnotics, agree not to alter the daily dose of these for the duration of the test.
5. If not on regular sedatives or hypnotics, agree not to commence these treatments during the test.
6. If on Parkinson's disease or psychotropic medication, doses stable for 1 month before and throughout the course of the study.
7. Able to provide informed consent.
8. Able to understand English.
9. Have a phone.
10. Agree not to drive or operate heavy machinery within 8 h of ingestion of study medication.

1.4.2. Exclusion criteria

1. Scores < 28 in Telephone Interview Cognitive Status (TICS) (Brandt J et al., 1988)
2. Has diagnosed sleep apnea or high risk of this (2 or more categories where the score is positive on Berlin questionnaire for sleep apnea risk [22]).
3. Co-morbid psychiatric/neurological diagnoses that may affect sleep, including:
   - Acquired brain injury
   - Uncontrolled major depression
   - Active or untreated post-traumatic stress disorder
   - Uncontrolled psychosis or schizophrenia
   - Unstable seizure disorder (i.e. seizure in the last 12 months)
   - Other relevant medical diseases, malignancy or other progressive neurological disorder
   - Cognitive impairment defined by Standardised Mini Mental State Score < 25/30 [23].
4. Known allergy or hypersensitivity to melatonin or previous adverse event from melatonin.
5. Contraindications to melatonin, such as on immunosuppressive drugs or anticoagulant drugs; patients with active or uncontrolled hormonal disorders, or diabetes, or significant active liver disease (determined by clinician), or moderate-severe abnormal kidney function (determined by clinician) or untreated kidney disease, or any blood clotting disorders.
6. Breastfeeding or pregnant women.

(Use of melatonin currently or in the last month is not an exclusion, but use needs to cease two days before the run-in period).

2. Study measures

2.1. Primary outcome measure

2.1.1. Parkinson's Disease Sleep Scale (PDSS-2)

The PDSS-2 assesses the subjective impact of aspects of sleep-related problems and insomnia in patients with PD [16]. The test-retest reliability (ICC) for a total score is 0.73. Example questions include – do you have difficulty falling asleep each night; do you have difficulty staying asleep? Scores range from 0 to 60, a higher score is worse, and the scores are treated as continuous.

The PDSS-2 cannot be used for screening. The PSQI is not suitable for outcome measurement. PSQI provides a comprehensible (to a broad audience) categorization of 'poor sleeper', the PDSS-2 is a PD specific, and appropriate, outcome indicator. PSQI provides some measure of severity of sleep disturbance that warrants treatment. The PDSS-2 is a better outcome measure as it takes into account additional benefits of improvement in sleep in PD population.

2.2. Secondary outcome measures

- Daytime sleepiness measured by NIH PROMIS sleep-related impairment scale [24].
- Mean change in SOL (min), measured by actigraphy and sleep diary. We will determine SOL by diary report of bedtime vs actigraphy algorithm to detect sleep onset.
- Sleep efficiency (proportion of time spent asleep while in bed) as measured by actigraphy.
- Adverse events, measured by the National Cancer Institute, US National Institutes of Health: Common Terminology Criteria for Adverse Events v5.0 [25].

Compliance will be monitored and reported.

2.3. Description of study measures

2.3.1. Pre-screening measures

2.3.1.1. Pittsburgh Sleep Quality Index (PSQI). The PSQI, a self-report measure of sleep quality, queries about multiple sleep-related variables over the preceding month, using Likert and open-ended response formats (Buysse D et al., 1989). A PSQI global score > 5 resulted in a sensitivity of 98.7 and specificity of 84.4 as a marker for sleep disturbances in insomnia patients versus controls.

2.3.1.2. Telephone interview for Cognitive Status (TICS). The Telephone Interview for Cognitive Status (TICS) is a global mental status test that can either be administered over the telephone or face-to-face [26]. Among elderly populations, TICS scores approximate a normal distribution and are not subject to the ceiling effects that limit the usefulness of many mental status examinations.

2.3.1.3. Berlin questionnaire. The Berlin questionnaire is a self-administered questionnaire that was developed to identify subjects with obstructive sleep apnea (OSA) in primary care settings [22]. A high-risk classification on the Berlin questionnaire in the primary care setting was associated with a sensitivity of 86% and specificity of 77% for a respiratory disturbance index (RDI) greater than 5 and a sensitivity of 54% and specificity of 97% for an RDI greater than 15, respectively.

The patient's neurologist will provide the Hoehn and Yahr scale score. The Hoehn and Yahr scale provides an estimate of clinical function in PD, focusing on motor symptom severity and relative level of disability (Hoehn M and Yahr M, 1967). Direct clinimetric testing of the HY scale has been very limited, but the scale fulfils at least some criteria for reliability and validity, especially for the midranges of the scale (Stages 2–4) [27].

2.4. Formal screening measures

2.4.1. Cognition: Standardised Mini-mental state examination (SMMSE)

The Mini–Mental State Examination is a 30-point questionnaire that is used extensively in clinical and research settings to measure cognitive impairmen and screen for dementiat [28]. The intra-rater variability is significantly lower with the SMMSE (86%, $p < 0.003$) and the inter-rater variance was reduced by 76%, compared to the MMSE. Intraclass correlation for the MMSE was 0.69 compared to 0.90 for the SMMSE.

2.5. Trial measures (baseline and 2 or 4 weekly thereafter)

2.5.1. Primary outcome measure - Parkinson's Disease Sleep Scale (PDSS-2)

See above for description.

2.5.2. Secondary outcome measures

2.5.2.1. NIH patient-reported outcomes measurement information system (PROMISTM) sleep-related impairment scale. [24] is a well-normed and psychometrically sound measure of the daytime impact of sleep difficulties. The scale is used to measure participants' self-reported tiredness, alertness, and sleepiness during normal waking hours.

2.5.2.2. Actigraphy. Objective measures of sleep parameters will be obtained from wrist-worn movement sensors supported by diary report, as per the American Academy of Sleep Medicine guidelines. We will use MotionWatch 8 (manufactured by Camntech). Objective determination of the following sleep parameters is possible from activity via published algorithms: sleep efficiency [time asleep/time in bed], sleep onset latency [time from 'lights out' to sleep onset], total sleep duration [time asleep to time of morning waking]. Sleep latency and sleep efficiency will be secondary outcome measures for statistical analysis. The other measures will go into the patient report to assist the patient and

clinician to make informed clinical decisions.

2.5.2.3. Adverse events. These will be measured by the National Cancer Institute, US National Institutes of Health: Common Terminology Criteria for Adverse Events v5.0 [25].

These criteria will be used to rate a reported event by severity.

2.5.3. Measures of covariates

2.5.3.1. Geriatric depression scale (GDS). The GDS-15 is a brief, 15-item questionnaire in which participants are asked to respond answering yes or no in reference to how they felt over the past week. In a validation study comparing the Long and Short Forms of the GDS for self-rating of symptoms of depression, both were successful in differentiating depressed from non-depressed adults with a high correlation ($r = 0.84$, $p < 0.001$) [29].

2.5.3.2. The Parkinson Anxiety Scale (PAS). The Parkinson Anxiety Scale is a 12-item patient-rated scale with three subscales, for persistent, episodic anxiety and avoidance behaviour [30]. It assesses the severity of anxiety in Parkinson's patients based on the extent and frequency of symptoms associated with the above disorders in the last 4 weeks. It is easy and brief to administer and has better clinimetric properties than existing anxiety rating scales.

2.5.4. Other measures

2.5.4.1. Patient recording of observations. Sleep diaries at the same time each day during run-in and trial phase (daily self-report of sleep parameters: time of 'lights out,' wake-up time, night-time activity; sleep medications). Diary data is also needed to define sleep periods for actigraphy.

2.5.4.2. REM Sleep Behaviour Disorder Single-Question Screen (RBD1Q). REM Sleep Behaviour Disorder Single-Question Screen (RBD1Q) is a screening question for dream enactment with a simple yes/no response. Compared to polysomnography to define gold-standard diagnosis according to standard criteria, there was a sensitivity of 93.8% and a specificity of 87.2%. A single-question screen for RBD may reliably detect disease, with psychometric properties favorably comparable to those reported for longer questionnaires [31].

2.5.4.3. Compliance with medication. Compliance will be measured by self-report and pill counts.

2.5.4.4. Feasibility. This will be assessed by N-of-1 test commencement and completion rates, withdrawal rates, and a post-test questionnaire for both doctors and participants, assessing their satisfaction with the N-of-1 test and changes in management of the patient's sleep problem.

2.6. Data collection

All forms will be available in online and paper copy formats. Table 1 illustrates the schedule and nature of data collection required during the study period.

2.6.1. Actigraphy

Objective measures of daily sleep-wake activity will be obtained with a small (36× 28.2 × 9 mm) and lightweight (9 g) triaxial accelerometer-based activity monitor. The monitor is worn on the non-dominant wrist continuously during the 14-day measurement periods. Accelerometry data will be processed in MotionWare proprietary software using published sleep identification algorithms validated for 30-s epochs. Standard sleep parameters (including sleep onset latency and sleep efficiency estimates) and non-parametric estimates of circadian rhythms parameters (including inter-daily circadian stability) [32] will be generated for each measurement period.

2.7. Participant recruitment and screening

The trial will be set up and co-ordinated by The University of Queensland coordination centre. Subjects will be recruited from Uniting Care Health physicians and through advertisements in the Parkinson's Queensland Incorporated website and Newsletter. Patients will see the research team at the clinical trial rooms at Wesley Medical Research, or St Andrew's War Memorial Hospital, Brisbane, or The University of Queensland Centre for Clinical Research.

2.7.1. Pre-screening, screening and registration

There are two possible routes for enrolment: Firstly, participating neurologists will discuss the trial briefly with patients who claim chronic sleep difficulty which is impacting their life. If they are interested the neurologist will send their contact details to the research nurse (RN). Alternatively, interested patients will self-refer from various sources.

The RN will contact the patient to have an initial discussion and telephone prescreening. Informed consent is obtained. The RN will contact their GP to do further medical prescreening using a prescreening form, which is also 1) a referral to the trial and 2) a trial "script" containing essential clinical information. If still potentially eligible, they will come in for a face to face appointment with the RN, during which the trial is explained in detail, and the rest of the eligibility criteria are checked. by the RN.

For screening, participants will complete the following:

- Pittsburgh Sleep Quality Index (PSQI) (Buysse D et al., 1989) (include if > 5)
- Telephone Interview Cognitive Screen (TICS) (Brandt J et al., 1988) (needs to be Unimpaired or Mild Impairment: exclude if $< 28/50$).
- Standardised Mini Mental State Score (SMMSS) face-to-face assessment by RN to exclude patients with dementia (exclude if $<$ 25/30 [23].
- A measure of the severity of Parkinson's Disease [33].
- Berlin questionnaire for sleep apnea risk [22] (exclude if 2 or more categories where the score is positive).

2.7.2. Obtaining informed consent

Written consent will be obtained from all participants.

2.8. Run-in period

There will be a two-week run-in period, on 3 mg melatonin. Standard scripts will cover dispensing a two-week supply of 3 mg melatonin to the patient. Sleep diaries are completed (including the PDSS-2 on the last day) during the two weeks on 3 mg. The diary is returned to the research staff, and the data is analysed. If the participant's PDSS-2 scores improve by at least 3.44 points, the test dose will be 3 mg; if the score does not improve, the test dose will be 6 mg. This assumes that the medication is well tolerated with no significant adverse events.

2.9. Randomisation

Once the test dose is determined the patient will be allocated a randomisation number. The script is sent to the study pharmacy.

Participants will undertake 3 pairs of treatment periods. In each pair, the order of treatment will be randomised. A randomisation list for the whole trial will be provided to the study pharmacy, who will randomise and pack the medications according to the schedule. This is also the study dispensing pharmacy, and they will then hold it for the duration of the study. The trial participants, investigators, clinicians and research staff will not know the treatment allocation. The randomisation schedule will consist of three separate randomisations of pairs of periods for all individuals in the study. The first period will be

Table 1
N-of-1 trial study schedule.

Evaluation	Pre-screening	Screening	Baseline	Run-in phase	Daily during trial	Every two weeks during trial	Just after trial	6 monthF/U	12 month F/U
Eligibility and informed consent	X	X							
Demographics	X	X							
Parkinson's Disease Severity:									
Modified Hoehn and Yahr			X						
Cognition:									
SMMSE ± TICS[a]	X	X							
Sleep:									
PSQI	X								
PDSS-2			X	X		X		X	X
Sleep Diaries				X	X			X	X
PROMIS sleep-related impairment			X			X			
RBD1Q	X		X			X			
Actigraphy				X	X				
Berlin questionnaire	X								
Health Status:									
Current Medications			X			X	X	X	X
Comorbidities			X						
Depression and Anxiety:									
GDS-15			X			X			
PAS (Self-report version) (4 weekly)			X			X			
Adverse Events (weekly)				X		X			
SAEs				X		X			
Compliance				X	X				
Satisfaction							X		

[a] TICS is a phone pre-screen for patients who are in the community at first contact with the RN.

randomised so that equal numbers receive melatonin and placebo in period 1. In order to balance potential confounders and to conceal allocation,we will randomise in blocks of 8, i.e all variations of AB, AB, AB.

2.10. Blinding

The trial will be triple-blinded. There will be individual unblinding for the health care provider and participant only, due to individual trial reports. Study personnel will not be unblinded until the end of the study. The placebo will have the same appearance, volume, weight, odour, and taste as active product.

2.11. Study drug

Participants will take either immediate release melatonin or placebo 30 min before bedtime. Immediate release melatonin was chosen because it signals sleep onset more strongly (with evidence that this can reduce sleep onset latency), while delayed release boosts overall melatonin profile (this may be better for reduction of night-time awakening). The immediate release formulation, as well as being more effective in producing the required effect on circadian rhythm, is less likely to worsen daytime somnolence as a potential side effect. Immediate release melatonin has also been used in previous studies in insomnia in PD. Immediate release melatonin will be sourced from Pharmanord, Denmark. Two doses will be available (3 mg or 6 mg), in personalised N-of-1 tests. These doses were chosen based on the previous studies of melatonin in PD mentioned above; practicalities; and clinical opinion. Each patient will trial the most appropriate dose for them, against placebo; for example, 3 mg vs placebo, or 6 mg vs placebo. This will be decided according to the results of the run-in period.

The medication kits will be randomised according to a randomisation schedule generated by the study's statistician. The pharmacy will dispense 6 blister packs numbered 1–6, corresponding to each fortnight's treatment.

2.12. Adherence to study medication

Adherence with the study medications will be assessed by: (1) weekly self-recorded medication intake, (2) counts of returned tablets following the completion of treatment and (3) the trial staff will ask questions about adherence during the planned telephone-based reviews starting at 1-week post-commencement. Participants will be asked to return all unused tablets for counting at the end of the treatment period in a reply-paid post satchel.

2.13. Concomitant treatments

Participants in both groups will be asked not to seek other treatments and where possible not to change current medications for the 12-week trial period. The nominated general practitioner will be notified in writing of the individual's participation in the trial. They will be asked within reason to refrain from referring or suggesting additional or alternative treatments to the individual for the trial 12 weeks after randomisation. We will request patients not to take sleeping pills as needed, but regular use of other hypnotic medication is allowed if the dose remains the same throughout the study. Similarly, other non-pharmacological interventions are permitted during the study if the treatment remains the same throughout the study.

2.14. Withdrawal of study participants

Participants can withdraw from the study at any time without prejudice to their current or future management. A participant must give verbal notification to study staff of their decision to withdraw. Participants will also be withdrawn if they develop any of the exclusion criteria.

2.14.1. Exit assessment

On completion of a participant's final assessment session, the Research Assistant will complete an exit questionnaire to document whether the participant completed the study or not and if not why not. Participants and their general practitioners will receive formal feed-

back on their N-of-1 trials, to inform further prescription of melatonin.

After consultation with their referring doctor about the results, the doctor will fill out a post-trial questionnaire asking about the post-trial management outcome and their satisfaction with the trial, and the patient will fill out a brief satisfaction questionnaire.

2.14.2. Early stopping

The study will be stopped if the literature indicates new findings especially in terms of serious adverse events, or if the Data Safety Monitoring Board (DSMB) decides that serious adverse events indicate that review of the study protocol is required. The protocol will be amended or the study stopped early if an excess of particular SAE/Serious Adverse Reactions/Serious Unexpected Suspected Adverse Reactions appear to be protocol related. Alternatively, the DSMB may recommend protocol review or changes without stopping the trial.

2.15. Long-term follow-ups

The study will continue for six- and twelve-months follow-up. At these times, a one-week sleep diary, data on any ongoing sleep medication and PDSS-2 will be collected. The data will be collected using electronic links to the Redcap database or by mailing the follow-up forms to participants.

3. Adverse events

Information about adverse effects of the medication will be sought from all participants during questioning at weekly intervals or via contact with trial staff at any time if participants have questions or concerns about the medication. If a participant reports an adverse event or side effect, medical investigators JOS and GM will call the participant to assess further and to decide upon what action should be taken. This may include dose alteration or withdrawal from the study if deemed necessary.

The most common side effects reported with melatonin usage are daytime sleepiness, headaches and dizziness. We have taken this into account with the dose run-in as tolerated. More severe side-effects are rare.

3.1. Detecting AEs and SAEs

The research team will ask about the occurrence of AEs/SAEs at least weekly during the study, and at 6 and 12-month follow-up. Participants will also be asked if they have been admitted to the hospital, had any accidents, used any new medicines or changed concomitant medication regimens.

3.2. Data Safety Monitoring Board (DSMB)

An independent DSMB will be constituted to evaluate the safety aspects of the study, involving experts in the field of study. A charter for the DSMB will be in place, and it will be responsible for safety reporting for the study. The DSMB consists of Dr Peter Allcroft, sleep physician, Assoc Prof Rob Ware, biostatistician and Dr Stephen Read, neurologist.

The DSMB has the responsibility to review all SARS to determine if a pattern is suspected of developing for a particular event. These events are usually so infrequent that it is very unlikely that significant differences will be demonstrated. The protocol will be amended or the study stopped early if the DSMB determines a pattern is developing for a particular SAR.

3.3. Overdose

Several cases of overdose have been reported post-marketing. Somnolence was the most reported adverse event. Most were mild to moderate in severity. Administration of daily doses of up to 300 mg of melatonin without causing clinically significant adverse reactions have been reported in the literature. If overdose occurs, drowsiness is to be expected. Clearance of the active substance is expected within 12 h after ingestion. No special treatment is required.

3.4. Unblinding procedures

The patients, clinicians, research staff and outcome assessors will be blinded. Clinical Principal Investigators will be able to unblind individual cases if the following criteria are met:

1. Emergency Unblinding - To make a clinical treatment decision or when an unexpected serious adverse event occurs.
2. During an unmasked analysis in accordance with the study analysis plan or at the requestion of the Data Safety Monitoring Board.
3. At the conclusion of the study to determine the effect of the intervention.

4. Statistical considerations

4.1. Sample size

Simulation of the primary outcome (PDSS-2) was used to explore the statistical power of this study for comparing placebo versus treatment, see Ref. [34]. For this study, 500 data sets were simulated, each with three cycles of treatment and placebo for each individual. It was assumed that 75% of individuals would complete at least one cycle, 65% of participants would yield data on all three cycles and 5% of participants would complete two cycles. These percentages were based on the results of [35] who conducted a series of N-of-1 tests in individuals with chronic neuropathic pain. (The other series of N-of-1 trials in PD were too small to provide meaningful completion data). For the linear multilevel random effects model used to simulate data (see below), the following prior information was obtained from Ref. [36]: Baseline PDSS-2: mean of 18.53 with a standard deviation of 10.09. We used a threshold of −3.44 points change in the PDSS-2 for detecting a clinically significant improvement, as recommended by Ref. [17]. A common variance was assumed for all outcomes from all cycles with sufficient washout to ensure no carryover effect. However, unfortunately no estimate of between-subject variability was available in the literature, so a conservative estimate was assumed such that intracluster correlation ($\sigma_b^2/(\sigma_b^2 + \sigma_w^2)$) remained at 50%, where σ_b^2 is the between-subject variance and σ_w^2 is the within-subject variance. The results of the simulation suggest that a sample size of 44 is large enough to yield at least a 90% chance of concluding there is a difference between treatment and placebo (at the 5% significance level) when the above difference exists. Due to the need to fit a large number of models, this exploration of statistical power was conducted within a frequentist framework in the R-package using the lme4 package.

Sample size for an equivalent RCT, for maintaining at least 90% power, is **364**, based on mean difference of 3.44, standard deviation of 10.09, significance level of 5% and two-sided *t*-test.

4.2. Data analysis

We will undertake 1) analysis of individual N-of-1 tests; and 2) aggregated analysis across individual N-of-1 studies using data from all periods in all participants.

N-of-1 test commencement, completion rates, and withdrawal rates will be reported. Data on randomised patients at baseline will be summarized using appropriate descriptive statistics.

Regarding patients on other hypnotic medications such as benzodiazepines, z-drugs or atypical antidepressants, we will conduct a stratified analysis comparing patients either naïve or exposed to concomitant psychotropic agents.

4.3. Aggregated analysis

Bayesian statistics combines prior information with information from the data (through the likelihood) to form a posterior distribution. It is this distribution on which all inferences are based, for example, an estimate of a treatment effect can be taken as the posterior mean, and the posterior standard deviation can be taken as the standard error of this estimate.

N-of-1 data will be combined using a Bayesian multilevel random effects model. This model will account for repeated measures on individuals over time and will return posterior estimates of overall treatment effect as well as the effect in each individual participant.

The following linear mixed effects model will be estimated for the primary outcome:

$$y_{ij} = \theta_{0i} + \theta_{1i}\tau_j + \varepsilon_{ij},$$

where $\theta_{0i} = \theta_0 + b_{0i}$, $\theta_{1i} = \theta_1 + b_{1i}$, (θ_0, θ_1) are the population parameters, y_{ij} is the primary outcome for treatment τ_j, $(b_{0i}, b_{1i}) \sim MVN(0, \Omega)$ are random effects based on variance-covariance matrix Ω and $\varepsilon_{ij} \sim N(0, \sigma^2)$ are residual error., for $i = 1, ..., N$ (the number of patients) and $j = 1, ..., n_i$ (the number of observations for patient i).

As the above model will be estimated within a Bayesian framework, prior distributions on the parameters should be specified. These will selected to be vague, and will be:

$$\theta_0, \theta_1 \sim N(0, 1e6), \quad \sigma \sim U(a, b), \quad \Omega^{-1} \sim W(\rho\Omega_\rho, \rho),$$

where 'N' denotes Normal, 'U' denotes Uniform and 'W' denotes Wishart, and a, b, Ω_ρ and ρ will be selected following the work of [37].

A variety of extensions to the above model will then be trialled to capture potential correlation of response through time and differences in variability of response to treatment. For example, different autoregressive and/or moving average models will be estimated to determine if the data are correlated through time and (if so) how such correlation should be modelled. This will include trialling different variance structures to determine whether patients respond to treatments differently in terms of the variability in each outcome. Further, different variables such as age, gender, depression, presence of RLS/RBD, hypnotic use etc will be included in the model to determine if the given outcome is related to such variables. And lastly, to determine if there is any evidence that a sufficient washout period was not included in the study, a term which accounts for potential carryover effects of the active treatment will be included. Undertaking this approach to model the trial data leads to substantial model uncertainty. Thus, a choice will need to be made about which model is most appropriate for the data. Once this choice is made, it will determine which extensions/additions to the model are significant (i.e. should be kept in the model) and which are not (i.e. can be removed). For this analysis, this choice will be made through formal Bayesian model selection procedures based on the posterior model probability (that is, the probability of a model given the data) with a preference for the model with the largest such probability.

Bayesian inference for all models will be conducted in WinBUGS with an estimate of the marginal likelihood (proportional to the posterior model probability) provided by the integrated nested Laplace approximation [38]. WinBUGS will be run for a total of 100,000 iterations for posterior sampling with the first 50,000 iterations being discarded as burn-in. These samples will then be thinned to remove autocorrelation. Convergence in the final samples will be checked using visual plots of chain history and the modified Gelman-Rubin statistics [39]. To ensure the results from fitting each model are robust to the prior specification (i.e. ensure priors are vague), a sensitivity analysis will be conducted. This will entail varying the values of the parameters in the specification of the prior to explore the effect these have on the results of the analysis, see Ref. [37] for an example.

Based on each model, the below summaries relating to response to melatonin will be presented for primary and secondary outcomes:

1. The posterior mean of the average difference between melatonin (combined 3 and 6 mg doses) and placebo for the entire patient cohort;
2. The associated 95% credible interval of the posterior mean given in (1); and
3. The posterior probability that the average difference between melatonin scores and placebo scores is at least the MCID (minimal clinically important difference), where available.

The above three summaries will also be presented at the individual level. Notably, (3) describes the likelihood that an individual patient will do better on the active treatment in future cycles.

We can formally identify individual responders based on the Bayesian model that is fitted to the whole data. Such a model provides individual (and population) estimates of the effectiveness of treatment, and hence can be used to determine which individuals experience substantial benefit while on treatment. Patient characteristics between responders and non-responders could then be compared to determine if there are any differences. Further, this comparison could also be formally undertaken within the Bayesian model when patient information can be assessed for usefulness in explaining between subject variability of random effects.

Secondary outcomes will be modelled as detailed above for the primary outcome. Adjustments will be made, as required, depending on the nature of the response variable (i.e. binary, count, etc).

From the post-trial questionnaire, data on doctor and participant satisfaction with the N-of-1 trial and its usefulness and practicality will be analysed using R, and frequencies will be reported.

4.4. Individual patient reports

For individual patient reports, individual effects of melatonin on PDSS-2 and safety will be reported, as will whether the effect was within the patient's expectation of significant improvement. We will also report:

- SOL (min) measured by sleep diaries and actigraphy
- Mean duration of sleep (min) while in bed, measured by sleep diaries and actigraphy.
- Sleep efficiency (proportion of time spent asleep while in bed), measured by actigraphy.

4.5. Quality assurance and data integrity

This trial will be conducted in accordance with the protocol, ICH Guidelines for Good Clinical Research Practice and within all relevant local ethical regulations. The integrity of trial data will be monitored by regularly scrutinising data sheets for omissions and errors.

A regulatory approved electronic data capture system (REDCap), with web hosting facility, will be used to collect all clinical and safety data. REDCap (Research Electronic Data Capture) is a secure, web-based application designed to support data capture for research studies, providing: 1) an intuitive interface for validated data entry; 2) audit trails for tracking data manipulation and export procedures; 3) automated export procedures for seamless data downloads to common statistical packages; and 4) procedures for importing data from external sources [40].

5. Discussion

Using N-of-1 tests to identify those PD patients who respond to melatonin would be clinically useful, novel and could potentially avoid adding an ineffective treatment which incurs an increased burden and cost.

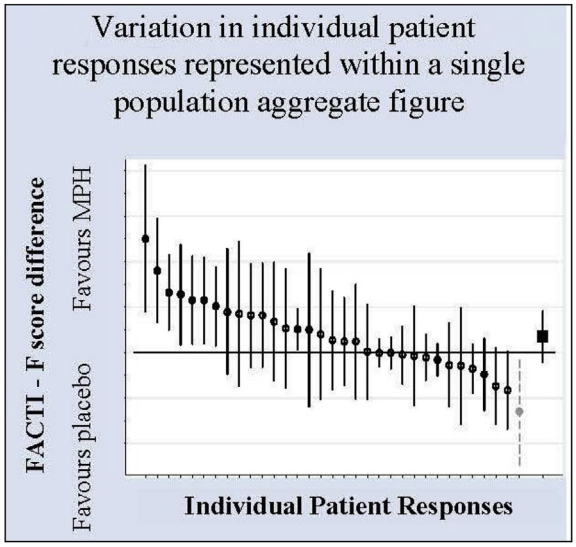

Fig. 2. Variation in individual patient responses represented within a single population aggregate figure.

Each participant will find out if melatonin helps their insomnia. For example, in another N-of-1 study, data obtained from 33 N-of-1 trials illustrates the variety of individual responses to methylphenidate (MPH), which was tested as a treatment for fatigue in palliative care patients (Fig. 2). There was no significant improvement when aggregated intervention data were compared to control overall, yet on an individual level, 8 patients were clear responders and one patient experienced greater fatigue on treatment (see Fig. 2) [14].

In this study, results from individual N-of-1 tests will be pooled to estimate the overall population treatment effect. The population estimate will show the likelihood that a person would be a responder, and thus the value of going through the N-of-1 process. Immediate release melatonin can be sourced from compounding pharmacies in Australia with a prescription.

This trial design will have no ability to detect potential long-term benefits (or adverse effects) of melatonin intake after 1 year.

Although, depending on the mechanism of insomnia, Cognitive Behaviour Therapy is current 'first line' treatment as recommended by peak bodies including the National Institutes of Health and the American Sleep Association, melatonin may become a useful alternative or adjunct therapy for insomnia in Parkinson's Disease. We anticipate that this trial may ultimately change the therapeutic options available to clinicians, PD sufferers and their families.

If n-of-1 tests for insomnia in PD are feasible and useful to patients and clinicians in this study, an N-of-1 trial, either in 3 × 2 cycles or in an abridged version, may be indicated to personalise the prescription of melatonin. There is the possibility of a further study to see if N-of-1 tests of melatonin in people with PD could be used in usual care.

6. Trial status

The study is currently recruiting.

Ethics approval

The study is approved by UnitingCare Health Research Ethics Committee #1702 and The University of Queensland Human Research Ethics Committee #2017000609.

Funding

The study is funded by Wesley Medical Research and The University of Queensland.

Conflict of interest

The authors have no conflict of interest to declare.

Acknowledgements

The authors thank Makram Bou-Ghannom, Shahiq Rizvi and Megdelawit Melaku for assistance with the study.

References

[1] Deloitte Access Economics, Living with Parkinson's Disease – Update, Accessed Jan 2018, 2011. https://www2.deloitte.com/au/en/pages/economics/articles/living-with-parkinsons-disease.html.
[2] P. Barone, et al., The PRIAMO study: a multicenter assessment of nonmotor symptoms and their impact on quality of life in Parkinson's disease, Mov. Disord. 24 (11) (2009) 1641–1649.
[3] L. Trotti, et al., Melatonin for sleep disorders in patients with Neurodegenerative diseases, Curr. Neurol. Neurosci. Rep. 16 (7) (2016).
[4] S. Mantovani, S.S. Smith, R. Gordon, J.D. O'sullivan, An overview of sleep and circadian dysfunction in Parkinson's disease, J. Sleep Res. 27 (3) (2018 Jun) e12673.
[5] L. Borek, J. Friedman, Mood and sleep in Parkinson's disease, J. Clin. Psychiatry 67 (6) (2006) 958–963.
[6] R. Romenets, et al., Doxepin and cognitive behavioural therapy for insomnia in patients with Parkinson's disease – a randomized study, Park. Relat. Disord. 19 (7) (2013) 670–675.
[7] K. Seppi, et al., The movement disorder society evidence-based medicine review update: treatments for the non-motor symptoms of Parkinson's disease, Mov. Disord. 26 (S3) (2011) S42–S80.
[8] W. Braam, et al., Exogenous melatonin for sleep problems in individuals with intellectual disability: a meta-analysis, Dev. Med. Child Neurol. 51 (2009) 340–349.
[9] A. Videnovic, et al., 'The clocks that time us': circadian rhythms in neurodegenerative disorders, Nat. Rev. Neural. 10 (12) (2014) 683–693.
[10] C. Garzón, et al., Effect of melatonin administration on sleep, behavioral disorders and hypnotic drug discontinuation in the elderly: a randomized, double-blind, placebo-controlled study, Aging Clin. Exp. Res. 21 (1) (2009) 38–42.
[11] W. Zhang, et al., Exogenous melatonin for sleep disorders in neurodegenerative diseases: a meta-analysis of randomized clinical trials, Neurol. Sci. 37 (1) (2016) 57–65.
[12] D. Breen, R. Barker, Exogenous Melatonin for Parkinson's Disease: "Waking up" to the Need for Further Trials, Park. Relat. Disord. (16) (2016) 30164–X 2016 May 13. pii: S1353-8020.
[13] D. Leger, et al., Nocturnal 6-sulfatoxymelatonin excretion in insomnia and its relation to the response to melatonin replacement therapy, Am. J. Med. 116 (2) (2004) 91–95.
[14] G. Mitchell, et al., The effect of Methylphenidate on fatigue in advanced cancer: an Aggregated -of-1 trial, J. Pain Symptom Manag. 50 (3) (2015) 289–296.
[15] L. De Lau, M. Breteler, Epidemiology of Parkinson's disease, B C Med J 5 (6) (2006) 525–535.
[16] C. Trenkwalder, et al., Parkinson's disease sleep scale – validation of the revised version PDSS-2, Mov. Disord. 26 (4) (2011) 644–652.
[17] K. Horvath, et al., Minimal clinically important difference on Parkinson's disease sleep scale 2nd version, Parkinson's Disease (2015), https://doi.org/10.1155/2015/970534.
[18] J. Ferreira, et al., Espresso coffee for treatment of somnolence in Parkinson's disease: results of n-of-1 trials, Front. Neurol. 7 (2016) 27.
[19] O. Rascol, et al., A proof-of-concept, randomized, placebo-controlled, multiple cross-overs (n-of-1) study of naftazone in Parkinson's disease, Fundam. Clin. Pharmacol. 26 (4) (2011) 557–564.
[20] A. Hughes, et al., Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases, J. Neurol. Neurosurg. Psychiatry 55 (3) (1992) 181–184.
[21] Daniel J. Buysse, Charles F. Reynolds, Timothy H. Monk, Susan R. Berman, David J. Kupfer, The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research, Psychiatr. Res. 28 (2) (1989) 193–213.
[22] N.C. Netzer, R.A. Stoohs, C.M. Netzer, K. Clark, K.P. Strohl, Using the Berlin Questionnaire to identify patients at risk for the sleep apnea syndrome, Ann. Intern. Med. 131 (7) (1999) 485–491.
[23] D. Molloy, T. Standish, A guide to the standardized Mini-Mental state examination, Int. Psychogeriatr. 9 (S1) (1997) 87–94.
[24] L. Yu, et al., Development of short forms from the PROMIS™ sleep disturbance and Sleep-Related Impairment item banks, Behav. Sleep Med. 10 (1) (2011) 6–24.
[25] US Department of Health and Human Services, Common Terminology Criteria for Adverse Events V5, 03 http://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf, (2015) Accessed 4/06/2016.
[26] J. Brandt, M. Spencer, M. Folstein, The telephone Interview for cognitive status, Neuropsychiatry Neuropsychol. Behav. Neurol. 1 (1988) 111-11.
[27] C.G. Goetz, W. Poewe, O. Rascol, et al., Movement disorder society task force report on the Hoehn and Yahr staging scale: status and recommendations, Mov. Disord. 19

(2004) 1020–1028.
[28] Molloy, et al., Reliability of a standardised MiniMental state examination compared with the traditional mini-mental state examination, Am. J. Psychiatry 14 (1991) 102–105.
[29] J.I. Sheikh, J.A. Yesavage, Geriatric Depression Scale (GDS): recent evidence and development of a shorter version, Clin. Gerontol. 5 (1/2) (1986 June) 165–173.
[30] A. Leentjens, et al., The Parkinson Anxiety Scale (PAS): development and validation of a new anxiety scale, Mov. Disord. 29 (2014) 1035–1043, https://doi.org/10.1002/mds.25919.
[31] R.B. Postuma, I. Arnulf, B. Hogl, A. Iranzo, T. Miyamoto, Y. Dauvilliers, W. Oertel, Y.E. Ju, M. Puligheddu, P. Jennum, A. Pelletier, C. Wolfson, S. Leu-Semenescu, B. Frauscher, M. Miyamoto, V. Cochen De Cock, M.M. Unger, K. Stiasny-Kolster, M.L. Fantini, J.Y. Montplaisir, A single-question screen for rapid eye movement sleep behavior disorder: a multicenter validation study, Mov. Disord. 27 (7) (2012 Jun) 913–916.
[32] D.L. Whitehead, A.D. Davies, J.R. Playfer, C.J. Turnbull, Circadian rest-activity rhythm is altered in Parkinson's disease patients with hallucinations, Mov. Disord. 23 (8) (2008 Jun 15) 1137–1145.
[33] M. Hoehn, M. Yahr, Parkinsonism: onset, progression and mortality, Neurology 17 (5) (1967) 427–442.
[34] K. Mengersen, J.M. McGree, C. Schmid, Statistical analysis of n-of-1 trials, in: J. Nikles, G. Mitchell (Eds.), The Essential Guide to N-Of-1 Trials in Health, Springer, Netherlands, 2015, pp. 135–153.
[35] M. Yelland, et al., N-of-1 Randomised trials to assess efficacy of Gabapentin for chronic neuropathic pain, Pain Med. 10 (4) (2009) 754–761.
[36] K. Horvath, et al., Test-retest validity of Parkinson's disease sleep scale 2nd version (PDSS-2), J. Parkinson's Dis. 4 (4) (2014) 687–691.
[37] C. Dansirikul, R. Morris, S. Tett, S. Duffull, A Bayesian approach for population pharmacokinetic modelling of sirolimus, Br. J. Clin. Pharmacol. 62 (2005) 420–434.
[38] H. Rue, S. Martino, N. Chopin, Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations, J. R. Stat. Soc. Ser. B 71 (2) (2009) 319–392.
[39] A. Gelman, D.B. Rubin, Inference from iterative simulation using multiple sequences, Stat. Sci. 7 (1992) 457–472.
[40] P. Harris, R. Taylor, R. Thielke, J. Payne, N. Gonzalez, J. Conde, Research electronic data capture (REDCap) - a metadata-driven methodology and workflow process for providing translational research informatics support, J. Biomed. Inform. 42 (2) (2009 Apr) 377–381.