Study Protocol and Statistical Analysis Plan

Project Title:
Impact of Hyperarousal on Simple and Complex Cognitive Task Performance Among Insomnia Sufferers

NCT Number: NCT02290405

Document Date: 18 May 2015

Most Recent IRB Approval: 02 Oct 2018

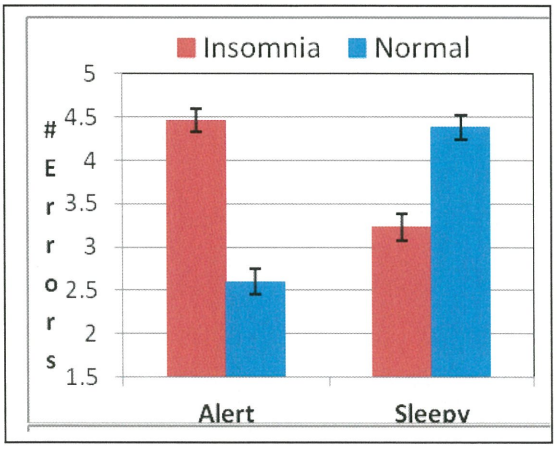
A

Background and Rationale

- Provide background on unanswered question(s) the study is attempting to answer (do not exceed one page)

Primary insomnia (PI) sufferers typically complain of such daytime impairments as reduced attention, concentration, memory and global mental acuity. Moreover, epidemiological studies have shown PI contributes to reduced productivity, work and traffic accidents, and serious falls among the elderly[1-3]. Despite such findings, laboratory-based efforts to corroborate the cognitive complaints of PI sufferers have produced mixed results. Indeed, many studies comparing PI sufferers with non-complaining normal sleepers across a range of neuropsychological tests have failed to show any relative deficits among the PI group[4,5]. Such findings, in turn, has led to the impression that PI patients cognitive complaints may be over-stated and result from their attentional bias toward minor cognitive errors, dysfunctional beliefs about the impact of insomnia on functioning[6] or excessive self focus[7] rather than to any measurable daytime impairment.

However, many previous such studies were underpowered due to small sample sizes and employed neuropsychological tests designed for detecting impairment resulting from brain disease/damage rather than the more subtle albeit significant impairments of which PI patients complain. In recent research, we[8] and others[9] have shown that PI sufferers do, indeed, show greater deficits (slower and more variable reaction times) particularly on complex switching attention tasks. Moreover, there is some preliminary evidence that the subgroup of PI sufferers with elevated levels of physiological hyperarousal are most prone to suffer from neuro-cognitive performance deficits than are matched groups of PI sufferers who are not physiologically hyperaroused and normally alert individuals without insomnia. For example, Fernandez-Mendoza[9] recently showed that PI sufferers with a hyperarousal pattern suggested by their objective short sleep duration on serial polysomnograms (PSG) performed more poorly on a complex switching attention task than did both normal sleepers and PI sufferers with normal objective sleep durations.



In our efforts to follow up on this latter work, we recently examined the error rates of alert and sleepy PI sufferers and normal sleepers across a series of simple and complex reaction time tasks. We employed age and gender matched samples of PI (N=89) sufferers and normal sleepers-NS (N=95). Participants underwent three nights of PSG followed by daytime testing with a four-trial Multiple Sleep Latency Test-MSLT. The PI and NS groups were each subdivided into "alert" (e.g., MSLT mean onset latency > 8 minutes) and "sleepy" (e.g., MSLT mean onset latency ≤ 8 minutes) subgroups to allow for testing the main and interaction effects of participant type and level of alertness. "Alert" participants had longer MSLT latencies than "sleepy" participants (12.7 vs. 5.4 minutes). PI sufferers had fewer correct responses on performance testing than did NS. However, as shown by the adjacent, figure we found a significant group x alertness interaction ($p = .0013$) with greater error rates occurring among alert (hyperaroused) PI sufferers (Mean=4.5±3.6 errors per trial) than among alert NS (Mean=2.6±1.9 errors per trial). This was particularly true for the more complex switching attention task.

Our work along with that of Fernandez-Mendoza[9] serve to confirm that PI sufferers have measureable objective neuro-cognitive deficits and provide some preliminary suggestion for the types of testing approaches that should be used to detect them. The identification of tests sensitive to PI sufferers' cognitive deficits are particularly relevant for testing the effects of current and future insomnia therapies on patients' objective daytime functioning. Measures of daytime dysfunction can and should serve as endpoints for assessing benefits and detriments of insomnia therapies. In addition, our recent work suggests that subgroups of PI sufferers may differ in their daytime deficits, with those showing physiological hyperarousal being most prone to make errors. This finding suggests that different types or doses of treatment may be needed to reverse the daytime impairments of the hyperaroused and non-aroused PI patients. However, our line of research would benefit by replication and extension findings to (1) further confirm the detrimental effects of physiological hyperarousal on PI sufferer's neuro-cognitive functioning; and (2) identify a broader range of tests that can be used for assessing diurnal cognitive impairments in both physiologically hyperaroused and lesser aroused PI groups. The current project will address these aims.

A

Objectives

- List the objectives to correspond directly with the listed hypotheses:

Aim 1: To compare the performances of hyperaroused PI sufferers and normally alert normal sleepers across a battery of computer-administered neuro-cognitive reaction time tests.

Aim 2: To examine the relative sensitivities of a range of simple and complex reaction time tests for discriminating PI sufferers from non-complaining normal sleepers.

Aim 3 (Exploratory): To examine the degree to which measures of the macro- and micro-architecture of sleep account for the daytime performance deficits shown by hyperaroused PI sufferers

Hypothesis

- List the clinical Hypotheses in order of priority:

Hypothesis 1: Hyperaroused PI sufferers will show greater performance deficits (slower reaction times; fewer correct responses and more error responses) across repeated daytime testing trials than will age- and gender-matched normal sleepers.

Hypothesis 2: Complex reaction time tasks that require executive function, ability to switch attention, and higher order memory function will show greater discrimination of hyperaroused PI sufferers and normal sleepers on measures of mean reaction time, correct response rates and error rates than will simple reaction time tasks.

Exploratory Aim 3: To determine the possible sleep deficiencies/mechanisms that contribute to the relative daytime deficits in the hyperaroused PI subgroups we will test measures of sleep continuity (sleep and wake time) and the macro-architecture (sleep stage percentages) and micro-architecture (NREM sleep EEG spectral measures) derived from PSG on nights preceding daytime testing as covariates in study analyses to determine the degree to which they account for observed group differences.

Study Design/Clinical Plan

- Provide a concise overview stating the type of experimental design

This study will use a matched-groups cross-sectional experimental design. Age and gender matched groups of hyperaroused PI sufferers and non-complaining normal sleepers (NS) will be recruited and enrolled. A comprehensive screening process that includes structured sleep and psychiatric interviews, screening questionnaires, vitals, urine pregnancy test, medical history, and diagnostic PSG will be used to determine eligible subjects. Our screening procedures and selection criteria are designed to allow us to identify and enroll hyperaroused PI sufferers and normally alert normal sleepers. PI sufferers enrolled will meet Research Diagnostic Criteria for insomnia disorder, score > 14 on the Insomnia Severity Index, report insomnia for > 3 months, have sleep difficulties $\geq$ 3 nights per week, score $\leq$ 3 on the Epworth Sleepiness Scale (ESS), score $\geq$ 29 on the Hyperarousal Scale[10] and report an inability to nap in the daytime. The normal sleepers enrolled will report general satisfaction with sleep and no sleep/wake complaints, score $\leq$ 10 on the ESS, score < 25 on the Hyperarousal Scale[10], and deny a practice of routine daytime napping. Those excluded will have: (a) a sleep-disruptive medical condition (e.g., rheumatoid arthritis); (b) a current major psychiatric (Axis I) condition on the basis of a Structured Clinical Interview for Psychiatric Disorders (SCID)[11]; (c) sedative hypnotic dependence and unwillingness/inability to abstain from these medications while in the study; (d) use of anxiolytics, antidepressants, or any other psychotropic medication; or (e) an apnea/hypopnea index (AHI) > 5 or a periodic limb movement-related arousal index > 15 during on screening PSG that includes a full sleep montage to allow for detection/diagnosis of sleep-disordered breathing and PLMD. Additionally, self-described NS who meet criteria for any sleep disorder and those insomnia sufferers who meet criteria for a comorbid sleep disorder in addition to PI will also be excluded.

Once consented, all participants will undergo screening procedures mentioned above. Those passing screening procedures will be scheduled for two consecutive nights of home-based PSG followed immediately by a daytime testing protocol. The nighttime PSG montage will include bilateral central and occipital referential EEG channels (C_4 and O_1, each referenced to $M_1 + M_2$), electro-oculogram (EOG, referenced to contralateral mastoid), submentalis electromyogram (EMG), and EKG. The PSGs will be used to derive sleep measures to address our exploratory Aim 3.

The daytime protocol will include a 4-trial Multiple Sleep Latency Test (MSLT) along with 4-trials of a computer –

administered battery of reaction time tasks. The assessment protocol will start two to three hours after participants' respective morning rising times and will begin with a battery of the neuro-cognitive testing followed by an MSLT nap. Per standard MSLT procedures, the daytime testing will be scheduled so the four performance testing and sleepiness assessment trials occur two hours apart. All daytime testing will be conducted under the supervision of trained laboratory technologists. Immediately prior to initiating daytime testing, participants' electrodes worn for the previous night's PSG will be checked and readjusted if necessary. Participants will be supervised between trials to prevent unscheduled sleep episodes. PSG electrodes will be worn for the entire day of laboratory testing and will not be removed until after the final trial is completed. Once the fourth performance/MSLT trial is completed and electrodes are removed, the participant will be allowed to leave the laboratory.

The neuro-cognitive test battery will include a 38 minute battery of simple and complex tests selected from CANTAB® Tests made available from Cambridge Cognition. The simple and complex tests will be presented in alternating order to control for the effects of fatigue on subjects' test performances. Among the simple tasks will be the Simple Reaction Time Test, Big Circle/Little Circle Test and Choice Reaction Time Test. These tests all provide measures of reaction time and motor speed but involve minimal cognitive load. The more complex tests will include the Rapid Visual Information Processing Test, The Spatial Working Memory Test and the Attention Switching Test. These latter tests collectively assess reaction time, attention, concentration, memory, and executive function, and thus, involve more cognitive load making them more sensitive to the deficits PI patients manifest. Both the simple and complex tests provide measures of respondents' mean reaction times, variability in reaction times across stimulus presentations, counts of correct responses, and counts of errors of omission and commission. These data are provided for each respondent for each trial allowing for analyses of respondents performances on a trial x trial basis as well as their average performances across trials.

Following each administration of the performance testing, the participant will be placed in a laboratory bedroom for the MSLT trial. Most aspects of the standard protocol[12] will be followed in conducting the MSLT. However, conservative MSLT criteria rather than contemporary clinical criteria will be used to define the sleep latency for each nap. Specifically, sleep latency will be defined as the time between the beginning of the nap trial and either the first three consecutive 30-second epochs of stage 1 sleep or the first 30-second epoch of any other sleep stage. If no sleep occurs, the trial will end at 20 minutes and a sleep latency of 20 minutes will be assigned. To minimize carry-over effects from one nap to the next, each nap trial will be discontinued 5 minutes after the sleep onset criterion is met.

Once all participants have completed the study protocol we will compare our groups of PI and NS in regard to their mean sleep onset latencies shown on the four MSLT naps to assure the groups do not differ in regard to their observed levels of daytime alertness. The dependent measures we plan to use in study analyses include: (1) mean reaction time per trial for each test; (2) # of correct responses per trial for each test; (3) # of errors of commission per trial for each test; (4) # of errors of omission per trial for each test; and (5) total # of errors per trial for each test. Our analysis plan below outlines how these measures will be used to test our study aims and hypotheses.

Treatment

- List the clinical dosage/dosage form, route, and dose regimen:

This section is not applicable since the proposed investigation does not involve medication dosing or include animals.

Collateral Research

- Include biomarkers, PK, etc.

Not applicable.

Statistical Plans

- Include justification for clinical sample size and primary hypothesis testing:

As noted above, our main dependent variables of interest will include: (1) mean reaction time per trial for each test; (2) # of correct responses per trial for each test; (3) # of errors of commission per trial for each test; (4) # of errors of omission per trial for each test; and (5) total # of errors per trial for each test. We will use these data to test Hypotheses 1-2 and to address our exploratory Aim which will examine the relationship between various nocturnal sleep measures and daytime performance scores. Our hypotheses testing will be accomplished using a 2 (PI vs. NS) x 2 (simple vs. complex test)

factorial analysis. Although our statistical model could also could include "trials' as a factor, we have chosen to limit the complexity of our analyses by excluding this factor and using mean values of each of our performance measures across trials for each participant in our analyses. To further simplify our analyses we will compute mean values of our 5 dependent measures for the three simple tests and another set of mean values of these dependent measures for our three complex tests and use those composite means in all planned analyses.

Before conducting tests of the study hypotheses, we will conduct a number of preliminary 2 (PI vs. NS) x 2 (alert vs. sleepy) x 2 (simple vs. complex test) ANOVAs and cross tabs analyses to determine if the various "cells" of our study include participants who do not differ significantly in terms of their mean ages, years of education, gender compositions and racial/ethnic compositions. If study "cells" are found to differ significantly on any of these variables, those variables will be used as covariates in statistical models designed to test study hypotheses. If study cells do not differ in regard to these variables we will use unadjusted statistical models to test study hypotheses.

For study hypotheses 1 & 2, we will conduct 2 (PI vs. NS) x 2 (simple vs. complex test) MANOVAs/MANCOVAs analyses to test for statistical significance across the multiple performance measures derived from the neuro-cognitive test battery. If the MANCOVA/MANOVA shows significant results, univariate ANCOVAs/ANOVAs will be conducted to tests for similar significant effects for each individual performance measure (i.e., mean reaction time; within subject standard deviations for reaction time; error responses).

Hypothesis 1 will be tested via examination of the main effect of the *group factor* (PI vs. NS) in these analyses. Findings showing significantly slower and more variable reaction times, fewer correct responses and higher error rates in the hyperaroused PI group will support this hypothesis. Elimination of these statistical differences in repeat analyses that use hyperarousal scale scores as a covariate will confirm the role of hyperarousal in explaining the observed group differences.
Hypothesis 2 will be tested via examination of the interaction of the *group factor* (PI vs. NS) and *test difficulty factor* (simple vs. complex) in these analyses. This finding coupled with a posteriori comparisons showing that PI participants have significantly poorer performances than do NS only on the complex tasks will support this hypothesis. Again, elimination of these statistical differences in repeat analyses that use each participant's hyperarousal scale score as a covariate will confirm the role of hyperarousal in explaining the observed group differences on complex tasks.

Our Exploratory Aim will require us to derive a variety of measures from the PSGs conducted on the two nights preceding daytime testing. Following the AASM Scoring Manual guidelines we will score these PSG records and obtain measures of total sleep time, sleep onset latency, wake time after sleep onset, sleep efficiency, and percentages of stages N1, N2, N3, and R for each night. We also will de-artifact each night and then perform a Fast Fourier Transform and subsequent sleep EEG spectral analysis as per our previously published reports[13,14] to derive NREM absolute and relative spectral power measures in the delta, theta, alpha, sigma, beta and gamma bandwidths. Once these measures are derived for each night, we will compute mean values of these measures for the two PSG nights and then enter those into multivariate regression analyses to determine which measures predict each of the five performance variables used as dependent measures. Those found to be predictive of performance indices will be entered as covariates in the above factorial analyses to determine if any such measures account for the anticipated relative deficits shown by the insomnia group vis a vis the normal group. Since we will anticipate group differences only on the 5 performance measures derived from the set of complex tests, we will plan to test our sleep measures as covariates only in one-way group comparisons using response latency, correct response rate, and error rate data acquired from the complex tests as dependent measures in these analyses. We, thus, will conduct and 5 ANCOVA analyses in which the sleep measures are tested as covariates with each of the 5 performance measures obtained from the complex reaction time tests.

In addition to these analyses, we will conduct simple bi-variate correlational analyses as well as multi-variate regression analyses to examine the association between the EEG measures and subjects' scores on the hyperarousal scale. In regression analyses hyperarousal scores will be predicted by the various measures derived from conventional and spectral scoring of the PSG records.

These analyses could provide important new information about the relationship between nighttime sleep and daytime performance. Specifically these analyses could help identify specific aspects of the macro and micro-architecture of sleep that are markers for sleep-disruptive hyperarousal and produce impairment in daytime neuro-cognitive performance. These sleep measures therefore could be considered important targets for insomnia therapies to address in optimizing the types of neuro-cognitive endpoints studied herein. Such findings would seem of great interest to Merck and other

A

industries that market agents for treating insomnia.

Power Considerations:

We conducted power calculations for the study proposed using data from our previous two studies[8,15] that compared reaction time performances of PI and NS participants. From the archival data sets used for these studies we obtained estimates of mean response latencies (reaction times) correct response rates and error rates for PI and NS groups to estimate the sample size needed to test the 2 study hypotheses. All power calculations were conducted using the proc power procedures available in SAS, version 9.2.

Hypothesis 1 entails testing the simple main effect of group membership (PI vs. NS) on neuro-cognitive performance. As a proxy for global performances we used response latency, correct response rate and error rate data derived from our 2008 study[8] and our above referenced pilot study[14]. Those studies showed the following information in comparisons of our hyperaroused PI group and normally alert NS group.

Measure	PI group	NS group	Sample standard deviation
Mean response latency - msec	470.6	420.8	112.1
Mean # correct responses	75.7	77.5	3.74
Means # errors	4.46	2.60	3.45

Given these data and assuming α = .05, and 82 subjects in each of these two groups we would have a .93 power to detect the predicted group differences in their mean error rates across trials, .87 power to detect the predicted differences in mean correct response rates, and .81 power to detect the predicted differences in mean reaction time latencies.

Hypothesis 2 entails testing the interaction term showing that difficult reaction time tasks will show significant differences between PI and NS whereas simple tests will not. Basically this analysis ultimately will entail a posteriori contrasts that show PI vs. NS group differences on complex tasks but not simple tasks. Again we used response latency data obtained from our 2008 study[8] for conducting power analyses for this hypothesis. Those data showed the following mean group differences and group standard deviations for complex and simple tasks.

Measure	PI group	NS group	Sample standard deviation
Complex test response latency - msec	496.65	438.14	133.4
Simple test response latency - msec	247.9	226.4	66.8

Given these data and assuming α = .05, and 82 subjects in each of these two groups we would have a.80 power to detect the predicted group differences on complex tasks but only .54 power to detect group differences on the simple tasks.

On the basis of these calculations we will plan to enroll 82 hyperaroused PI sufferers and 82 NS in this study.

Budget Summary

- Please be sure to complete budget template (excel document)

Total Amount Requested: (Include overhead)	**$501,882.68**
Additional sources of funding required? (Yes/No) If Yes, please be specific.	**No**

Timelines and Study Plans

Number of Sites:	**1**
Site Names:	National Jewish Health
Study Start Date:	October 1, 2014

A

Study End Date:	January 31, 2017
Number of Subjects:	164
First Patient In Date:	January 5, 2015
Last Patient Out Date:	January 31, 2017
Enrollment Period in Months:	36
Publication Plan	
Where are you planning to submit for publication? (journals, etc):	**SLEEP**
Are you planning to present your data at a scientific meeting?	Yes
Please list your target date for submission of publication.	**1/1/2016**
Drug Supply Information	
Drug Supplies Required (Yes/No)?	**No**
List Drug Supplies and Amount Required:	Drug Name: Amount:
List Drug Supplies and Amount Required:	Drug Name: Amount:
Placebo Required (Yes/No)?	
Additional Sources of Drug Supply (Yes/No). If Yes, please specify	**No**