

MEMORANDUM

DATE: May 2, 2006

FROM: Director
Division of Neurology Products/HFD-120

TO: File, NDA 21-875

SUBJECT: Action Memo for NDA 21-875, for the use of Nuvigil (armodafinil) to treat the excessive sleepiness of Obstructive Sleep Apnea/Hypopnea Syndrome (OSAHS), Narcolepsy, and Shift Work Sleep Disorder (SWSD)

NDA 21-875, for the use of Nuvigil (armodafinil) to treat the excessive sleepiness of Obstructive Sleep Apnea/Hypopnea Syndrome (OSAHS), Narcolepsy, and Shift Work Sleep Disorder (SWSD), was submitted by Cephalon, Inc., on 3/31/05. The application contained reports of four randomized, controlled trials, 2 in OSAHS, and one each in narcolepsy and SWSD. In addition, the requisite CMC, non-clinical, pharmacokinetic, and safety data were submitted.

Armodafinil is the R-isomer of Modafinil, the racemate, which has been marketed for several years for the same indications that are now proposed for armodafinil in this application. Plasma levels of R-modafinil (armodafinil) achieved with a given dose of Nuvigil are half those achieved after the same dose of the racemate.

This application has been reviewed by Dr. Norman Hershkowitz, medical officer, Dr. Jialu Zhang, statistician, Dr. Ta-Chen Wu, Office of Clinical Pharmacology, Dr. David Claffey, chemist, Jeanine Best, Division of Surveillance, Research, and Communication Support, Dr. Laura L. Pincock, Division of Medication Errors and Technical Support, Carolanne Currier, Division of Scientific Investigations, Dr. Sriram Subramaniam, Division of Scientific Investigations, and Dr. Geoffrey Zeldes, Controlled Substance Staff. The clinical team recommends that the application be considered approvable. I will very briefly review the relevant findings, and offer the rationale for the division's action.

One important aspect of the review needs to be noted at this point.

As Drs. Hershkowitz and Zhang both point out, the review of this application was marked by Dr. Zhang's finding of a significant problem in the conduct and reporting of the study results.

Specifically, each study employed both an objective measure of sleepiness, and a subjective measure. In the OASHS and narcolepsy studies, the objective measure was the Maintenance of Wakefulness Test (MWT). In the SWSD study, the objective measure was the Multiple Sleep Latency Test (MSLT).

In the MWT, patients are instructed to lie quietly and attempt to stay awake. Sleep latency is defined as the time to the onset of the first of 3 consecutive epochs of Stage 1 sleep, or the time to onset of any epoch of Stage 2, 3, 4, or REM sleep, as assessed by polysomnography (PSG). The PSG tracings were divided into 30 second epochs, and patients were assigned a sleep stage if they were in a particular stage for at least 16 seconds of an epoch. If no sleep occurred, a sleep latency of 20 or 30 minutes was assigned (the sessions varied in duration). The assignment of sleep latency was made by technicians observing the patients. If the technician observed that the patients had fallen asleep and had met the criteria for assigning a specific sleep latency, the patient was awoken, and the PSG tracings were terminated. The MWT was performed 6 times (0900, 1100, 1300, 1500, 1700, and 1900) at Weeks 4, 8, and 12. Only the mean of the first 4 sessions was to be used as the primary data.

However, the PSG tracings were then read by a central reader, after the study. Dr. Zhang found that in many cases, the central reader disagreed with the determination of the technician on site; that is, in numerous cases in which the technician had determined that sleep had occurred, the central reader determined that the patient had not fallen asleep, and, in these cases, the central reader assigned the maximum 20 or 30 minutes of sleep latency. Because of these discrepancies and the resulting effect on the analyses, we asked the sponsor to determine how often this occurred, and to re-analyze the data using numerous imputation schemes. These analyses were as follows:

Analysis 1

Use the latency as assigned by the local technician.

Analysis 2

Use LOCF for endpoint visits with a flawed session. That is, if any of the 4 sessions was flawed, use the previous (Week 4 or 8) visit with no flawed sessions. If all of these 3 visits are flawed, the unflawed experimental median endpoint score for the combined drug/placebo group.

Analysis 3

For a session with a flawed nap, use the fifth (or if necessary, the fifth and sixth naps) of that session. If 4 unflawed naps do not exist for that session, use the unflawed median of the combined drug/placebo score for the relevant naps.

Analysis 4

Discard flawed sessions and replace with the median of unflawed scores for the combined drug/placebo score for the first 4 naps for that particular stage.

Analysis 5

Discard data and do not analyze patients with any of the first 4 naps/session at baseline or final endpoint.

Analysis 6

Analyze only those sessions in which the first 4 naps are unflawed.

Analysis 7 **Worst case scenarios**

Analysis 7.1

In the placebo group, replace each flawed baseline session with the lowest sleep latency of 6 baseline sessions for each corresponding patient and replace each flawed endpoint session with the highest sleep latency of unflawed sessions from the 6 endpoint sessions of that patient. In the drug group, assign the highest and lowest latencies, respectively.

Analysis 7.2

As above, but use only the first 4, not 6 sessions.

Analysis 8 **Most Conservative worst case analyses**

Analysis 8.1

In the placebo group, replace each flawed baseline session with the lowest sleep latency of the 6 baseline sessions for each patient (as above), and replace each flawed session at endpoint with the full duration (either 20 or 30 minutes). In the drug group, replace each flawed baseline session with the full duration, and replace each flawed endpoint session with the lowest latency of the 6 endpoint sessions.

Analysis 8.2

As above, but use only the first 4, not 6 sessions.

As Dr. Zhang describes, in the OASHS and narcolepsy studies, the number of flawed sessions varied from 6-15%, and the number of patients with at least one flawed session varied from 17-41%, depending upon assigned group. In the SWSD study, the number of flawed sessions varied from 0.6-1.5%, and the number of patients with at least one flawed session varied from 0-4.5%.

Study 3020

This was a 12 week, multi-center double-blind study in patients with narcolepsy who were randomized to receive Nuvigil 125 or 250 mg, given once a day, or placebo. The co-primary outcomes were the mean change from baseline in the sleep latency of the average of the first 4 naps of the MWT, and the proportion of patients with at least minimal improvement on the CGI-C.

As Dr. Zhang notes, all of the analyses through Analysis 6 reached statistical significance for the MWT for both doses, with most showing a slight, but definite, superiority of 250 mg dose compared to the 150 mg dose. Analyses 7.1 and 7.2 were also nominally significant for the 250 mg-placebo contrasts.

The analyses of the CGI-C were highly significant for both groups, with 73 and 69% of the 250 and 150 mg/day groups, respectively, showing some improvement.

Study 3021

This study was of similar design to that of 3020, but enrolled patients with OSAHS.

Here again, all analyses through Analysis 6 reached significance for both dose groups, with only very slight numerical superiority of the 250 mg group compared to the 150 mg group. Here, Analysis 7.1 yielded a p-value for the 250 mg vs placebo contrast of 0.06, and Analysis 7.2 yielded p-values for the 250 mg vs placebo and 150 mg vs placebo contrasts of 0.26 and 0.048, respectively.

The analyses of the CGI-C were highly significant, with 74 and 71% of the 250 mg and 150 mg groups, respectively, showing improvement.

Study 3025

This was study similar in design to Study 3021, but patients were randomized only to either Nuvigil 150 mg or placebo.

Here, all analyses through Analysis 6 reached statistical significance. The results of Analysis 7.1 yielded a p-value of 0.14, and the results of Analysis 7.2 yielded a p-value of 0.03.

The analysis of the CGI-C was highly significant ($p=0.007$).

Study 3022

This was of similar design to the previous studies, except in patients with SWSD, and the co-primary outcomes were the mean change from baseline in sleep

latency as assessed by the MSLT and the proportion of patients with improvement on the CGI-C. In this study, patients were randomized to either Nuvigil 150 mg/day or placebo.

All analyses reached statistical significance; the worst-case analyses were not performed, but would have been expected to have reached significance. The p-value for the analysis of the CGI-C was 0.001).

SAFETY

A total of 1271 patients with either OASHS (N=751), narcolepsy (N=233), or SWSD (N=287) have received at least one dose of armodafinil in either controlled or open-label trials. A total of 781 have received armodafinil for at least 6 months, and 335 have received the drug for at least one year.

One patient, a 59 year old man with OASHS died after about 6 months of treatment. He was found dead in bed several days after having been diagnosed with pneumonia. He was receiving concomitant rofecoxib for arthritis.

A total of 46 patients experienced at least one serious adverse event. The most common single adverse event that was considered serious was chest pain, reported in 5 patients (<1%). The only other single events that occurred in more than one patient (N=2 each) were myocardial infarction and hemorrhoidal hemorrhage.

In controlled trials, a total of 6 ADRs were considered serious in patients treated with armodafinil (0.9%), compared to 2 in the placebo group (0.4%). The serious ADRs in the armodafinil group that were possibly related to treatment were Migraine, Affective disorder and Personality disorder (in the same patient), and depression, although the latter two patients had histories of psychiatric disorders; a case of angioneurotic edema seemed likely related to drug.

A total of 132 patients (10.3%) withdrew from studies secondary to adverse events. In controlled trials, a total of 44 patients (7%) withdrew armodafinil treatment due to an adverse event, compared to 16 placebo patients (4%). No single adverse event that led to withdrawal occurred at a rate greater than 1%. The following list, in decreasing frequency, led to treatment withdrawal in more than a single patient in controlled trials: Headache (8), Anxiety, Depression, Nausea (4), Palpitations, Diarrhea, AST increased, GGT increased, Agitation, Insomnia (3), Cardiac Flutter, Blurred Vision, ALT increased, Dizziness, Migraine, Nervousness, Sleep Disorder, Dyspnea, Rash (2). Reasons for withdrawal in open label studies were similar to those that resulted in withdrawal in the controlled trials.

Common adverse events

Common adverse events that occurred at least 2% and were more frequent than in the placebo group are listed below:

Event	Nuvigil (%)	Placebo (%)
Headache	17	9
Nausea	7	3
Dizziness	5	2
Insomnia	5	1
Diarrhea	4	2
Dry Mouth	4	<1
Anxiety	4	<1
Depression	2	<1
Rash	2	<1
Fatigue	2	1
Dyspepsia	2	<1
Palpitations	2	1

Laboratory Measures

Overall, Nuvigil did not cause important changes in routine laboratory measures. However, there was a slight mean increase (compared to placebo) in serum GGT (difference from placebo in change from baseline of about 7.2 U/L; baseline mean about 34-39), and a slight decrease in mean uric acid (difference from placebo in change from baseline of about -14.5 micromol/L; baseline mean of between 290-390). These changes were likely clinically unimportant, and there were no obvious differences in the proportion of patients who met outlier criteria.

There were no important overall changes in laboratory measures, but there was one patient who developed mild pancytopenia after 35 days of treatment (see Dr. Hershkowitz's review, Laboratory results/hematology) which resolved after having been taken off drug for about 2 weeks. Also, 6 patients discontinued treatment for elevated ALT/AST (without bilirubin elevation) that also resolved off treatment.

Vital Signs/EKG

In general, Nuvigil was associated with small, but consistent, increases in mean systolic and diastolic blood pressure (mean difference from placebo in change from baseline of 1-4 mm Hg systolic and about 1-3 mm Hg for diastolic, with a suggestion of a dose response). There were also slight increases in the proportion of patients who met various outlier criteria for elevated blood pressure (see Dr. Hershkowitz's review, vital signs), as well as in the proportion of patients who initiated treatment with anti-hypertensive medications or who had

their doses of anti-hypertensives increased (see Dr. Hershkowitz's review, vital signs).

There were no important EKG changes.

COMMENTS

The sponsor has submitted the results of four randomized controlled trials that purport to establish the effectiveness of NUVIGIL in the treatment of the excessive sleepiness of OSAHS, narcolepsy, and SWSD. As noted, the conduct of these studies raised questions about the interpretability of the results. Because of this, we asked the sponsor to perform multiple additional analyses. In my view, these analyses establish the effectiveness of NUVIGIL for these indications. It is true, as the clinical team notes, that the results of the most conservative worst-case analyses fail to reach statistical significance. However, as Dr. Hershkowitz notes, these analyses were likely considerably too conservative, and, as he also notes, several of the other, but still relatively conservative worst-case analyses, did reach statistical significance. In my view, the results on these latter analyses, as well as on all of the other additional requested analyses, clearly establish NUVIGIL's effectiveness.

There are no safety issues that would preclude approval; what issues have been described above can be described in labeling. However, there are two additional clinical safety issues that require discussion.

Cephalon, Inc. has submitted to the Division of Psychiatry Products (DPP; HFD-130) an application for Modafinil for the treatment of pediatric patients with Attention Deficit/Hyperactivity Disorder (ADHD). In that dataset, a single presumptive case of Stevens-Johnson Syndrome (SJS) was noted. As a result, the Modafinil for ADHD application was discussed at a recent meeting of the Psychiatry Drugs Advisory Committee (PDAC). That committee recommended that the application not be approved until and unless an additional 3000 pediatric patients were exposed with no additional cases of SJS found.

Within the last several weeks, however, the sponsor has submitted to DPP photographs of the patient who experienced the rash; the sponsor believes that these photographs establish that the patient did not have SJS. These photographs are currently being reviewed by DPP and its consultants.

Although the sponsor has not proposed that NUVIGIL be approved for pediatric patients, the clinical team believes that, if modafinil is associated with SJS in pediatric patients, that fact should be prominently included in product labeling for NUVIGIL. I agree. However, such a statement (if appropriate) cannot be written until the issue of SJS and Modafinil has been definitively addressed. For this reason, until DPP has adequately evaluated this new information, final labeling for NUVIGIL cannot be written. In this regard, it will also be important for the

sponsor to go back to the Modafinil clinical trials database and undertake a detailed search for additional cases of serious rash (we are not confident, at this point, that such a search has been done). We are aware of several post-marketing reports of SJS in patients being treated with Modafinil, but the reporting rate is below the background rate.

Also, Dr. Hershkowitz has just received a post-marketing report of a 39 year old man who died of cardiac failure presumably related to multi-organ failure related to a hypersensitivity reaction to recently started Modafinil. This patient had also been treated with Trileptal (known to be associated with such reactions), but had been treated with Trileptal for many years. We will ask the sponsor for more information on this case, and also ask them to further evaluate their Modafinil database for similar reactions.

The pharmacology review staff notes several significant deficiencies in some of the non-clinical studies. Although we agreed with the sponsor that a very truncated non-clinical program would suffice as support for the NUVIGIL application (because of available data for modafinil to which NUVIGIL was linked), closer examination of several of the nonclinical studies done with modafinil reveals that they were inadequate. Specifically, modafinil has not had an adequate carcinogenicity evaluation. A study accepted as being adequate at the time for modafinil (by the CAC and the division, and later by the division as being adequate for armodafinil) was a Tg.AC mouse alternative assay, in which drug is delivered dermally. There is general agreement at this time, however, that this assay is entirely inadequate for drugs given orally, and, specifically in this case, we are unsure that this species is exposed to relevant circulating moieties that are present when armodafinil is given orally. Although, as noted, the division had agreed that the modafinil nonclinical studies would be adequate for armodafinil, upon further reflection, we now realize that this assay is inadequate; we will ask the sponsor to repeat an appropriate assay in Phase 4.

Similarly, one of the pre-post natal studies previously performed was inadequate (see Dr. Freed's memo).

b(4) b(5)

Finally, DMETS has determined that the similarity between Nuvigil 150 mg and Norinyl 1/50 when written is so great that errors are likely to occur, and that, therefore, the name NUVIGIL should not be permitted. I have discussed this with the team. I believe that the marked differences in packaging between the products (and the public's general knowledge of the unique packaging for oral contraceptives), as well as the fact that Norinyl is very rarely prescribed, mitigate these concerns. However, we will ask the sponsor to report any real or potential medication errors as 15 day reports, and to include the possibility of this error in its promotional materials.

ACTION

The sponsor has provided substantial evidence of effectiveness for NUVIGIL for its proposed claims. However, several issues, both clinical and nonclinical, and described above, need to be resolved prior to its approval. For these reasons, I will issue an Approvable letter with attached labeling.

Russell Katz, M.D.

Appears This Way
On Original

**This is a representation of an electronic record that was signed electronically and
this page is the manifestation of the electronic signature.**

/s/

Russell Katz
5/3/2006 07:40:43 AM
MEDICAL OFFICER

Appears This Way
On Original

MEMORANDUM
DEPARTMENT OF HEALTH AND HUMAN SERVICES
FOOD AND DRUG ADMINISTRATION
CENTER FOR DRUG EVALUATION AND RESEARCH
CONTROLLED SUBSTANCE STAFF

Date: April 5, 2006

To: Russell Katz, M.D., Director
Division of Neurology Products (HFD-120)

Through: Deborah Leiderman, M.D., Director
Michael Klein, Ph.D., Team Leader
Controlled Substance Staff (HFD-009)

From: Geoffrey Zeldes, M.D., Pharm.D., Medical Officer
Controlled Substance Staff (HFD-009)

Subject: CSS Abuse Liability Assessment of NDA 21-875, Nuvigil
(armodafinil) oral tablets (50 mg, 100 mg, 150 mg, 250 mg IR tablets)
Indication: treatment of excessive sleepiness with obstructive sleep
apnea, hypopnea syndrome, narcolepsy, and shift work sleep disorder
Date of Submission: April 13, 2005
Sponsor: Cephalon

Background

Provigil[®] (modafinil) is approved by the FDA for the treatment of excessive sleepiness associated with narcolepsy, obstructive sleep apnea/hypopnea syndrome and shift work sleep disorder. The sponsor intends to market Nuvigil (armodafinil) for similar indications.

Armodafinil is the R-enantiomer of modafinil. Modafinil is a racemic mixture of R and S enantiomers. The Sponsor claims that each enantiomer displays similar CNS receptor site and pharmacological activity in a similar dose range.

Armodafinil exhibits linear, time independent pharmacokinetics following oral single- and multiple- dose administration. T_{max} is about 2 hrs in the fasted state; absorption is delayed about 2 -4 hours by food. The elimination half life of armodafinil is about 15 hours. The armodafinil (the R-enantiomer) half life in humans is approximately three times that of the S-enantiomer as a result of the clearance of the S-enantiomer being three-fold faster than that of the R-enantiomer. The Sponsor states that studies show that the enantiomers do not interconvert. At steady state of modafinil, total exposure to armodafinil is approximately three times that for the S-isomer.

Cognitive Drug Research (CDR) system, and behavioral attention measured by the Psychomotor Vigilance Task (PVT).

The Division of Scientific Investigations has evaluated Study _____ and expressed concerns regarding using a _____ assay to determine the concentration of R-modafinil in human plasma. _____

b(4)

Appears This Way
On Original

**This is a representation of an electronic record that was signed electronically and
this page is the manifestation of the electronic signature.**

/s/

Geoffrey Zeldes
4/5/2006 12:37:05 PM
MEDICAL OFFICER

Michael Klein
4/5/2006 12:48:30 PM
CHEMIST
Acting Director - for Deborah Leiderman, MD, Director, CSS

Appears This Way
On Original

CLINICAL REVIEW

Application Type NDA
Submission Number 21,875 (000)

Letter Date 3/31/05
Stamp Date 3/31/05
PDUFA Goal Date 4/30/06

Reviewer Name Norman Hershkowitz MD, PhD
Review Completion Date 4/30/06

Established Name Armodafinil
(Proposed) Trade Name Nuvigil
Therapeutic Class Wakefulness Promoting Agent
Applicant Cephalon

Priority Designation S

Formulation Tablets (50, 100, 150 and 250 mg)
Dosing Regimen 150 and 250 qD
Indication Improve Wakefulness
Intended Population OSAHS, SWSD and Narcolepsy

Table of Contents

1. EXECUTIVE SUMMARY	5
1.1 RECOMMENDATION ON REGULATORY ACTION	5
1.2 RECOMMENDATION ON POSTMARKETING ACTIONS	6
1.2.1 Risk Management Activity	6
1.2.2 Required Phase 4 Commitments	6
1.2.3 Other Phase 4 Requests	6
1.3 SUMMARY OF CLINICAL FINDINGS	6
1.3.1 Brief Overview of Clinical Program	6
1.3.2 Efficacy	7
1.3.3 Safety	10
1.3.4 Dosing Regimen and Administration	13
1.3.5 Drug-Drug Interactions	14
2. INTRODUCTION AND BACKGROUND	15
1.4 PRODUCT INFORMATION	15
1.5 CURRENTLY AVAILABLE TREATMENT FOR INDICATIONS	15
1.6 AVAILABILITY OF PROPOSED ACTIVE INGREDIENT IN THE UNITED STATES	15
1.7 IMPORTANT ISSUES WITH PHARMACOLOGICALLY RELATED PRODUCTS	16
1.8 PRESUBMISSION REGULATORY ACTIVITY	17
OTHER RELEVANT BACKGROUND INFORMATION	17
3. SIGNIFICANT FINDINGS FROM OTHER REVIEW DISCIPLINES	17
1.9 CMC (AND PRODUCT MICROBIOLOGY, IF APPLICABLE)	17
1.10 ANIMAL PHARMACOLOGY/TOXICOLOGY	17
4. DATA SOURCES, REVIEW STRATEGY, AND DATA INTEGRITY	18
1.11 SOURCES OF CLINICAL DATA	18
1.12 TABLES OF CLINICAL STUDIES	19
1.13 REVIEW STRATEGY	19
1.14 DATA QUALITY AND INTEGRITY	19
1.15 COMPLIANCE WITH GOOD CLINICAL PRACTICES	20
1.16 FINANCIAL DISCLOSURES	20
5. CLINICAL PHARMACOLOGY	20
6. INTEGRATED REVIEW OF EFFICACY	20
1.17 INDICATION	20
1.17.1 Methods	20
1.17.2 General Discussion of Endpoints	21
1.17.3 Study Design	29
1.17.4 Efficacy Findings	30
1.17.5 Clinical Microbiology	67
1.17.6 Efficacy Conclusions	67
7. INTEGRATED REVIEW OF SAFETY	70
1.18 METHODS AND FINDINGS	70
1.18.1 Deaths	70
1.18.2 Other Serious Adverse Events	70
1.18.3 Dropouts and Other Significant Adverse Events	76
1.18.4 Other Search Strategies	92

1.18.5	Common Adverse Events.....	92
1.18.6	Less Common Adverse Events	99
1.18.7	Laboratory Findings.....	99
1.18.8	Vital Signs.....	111
1.18.9	Electrocardiograms (ECGs).....	115
1.18.10	Immunogenicity	118
1.18.11	Human Carcinogenicity	118
1.18.12	Special Safety Studies.....	119
1.18.13	Withdrawal Phenomena and/or Abuse Potential.....	121
1.18.14	Human Reproduction and Pregnancy Data	121
1.18.15	Assessment of Effect on Growth.....	121
1.18.16	Overdose Experience	121
1.18.17	Postmarketing Experience.....	121
1.19	ADEQUACY OF PATIENT EXPOSURE AND SAFETY ASSESSMENTS	122
1.19.1	Description of Primary Clinical Data Sources (Populations Exposed and Extent of Exposure) Used to Evaluate Safety	122
1.19.2	Description of Secondary Clinical Data Sources Used to Evaluate Safety	126
1.19.3	Adequacy of Overall Clinical Experience.....	126
1.19.4	Adequacy of Special Animal and/or In Vitro Testing.....	127
1.19.5	Adequacy of Routine Clinical Testing.....	127
1.19.6	Adequacy of Metabolic, Clearance, and Interaction Workup	127
1.19.7	Adequacy of Evaluation for Potential Adverse Events for Any New Drug and Particularly for Drugs in the Class Represented by the New Drug; Recommendations for Further Study	127
1.19.8	Assessment of Quality and Completeness of Data.....	127
1.19.9	Additional Submissions, Including Safety Update.....	127
1.20	SUMMARY OF SELECTED DRUG-RELATED ADVERSE EVENTS, IMPORTANT LIMITATIONS OF DATA, AND CONCLUSIONS	136
1.21	GENERAL METHODOLOGY	136
1.21.1	Pooling Data Across Studies to Estimate and Compare Incidence	136
1.21.2	Causality Determination	140
8.	ADDITIONAL CLINICAL ISSUES.....	141
1.22	DOSING REGIMEN AND ADMINISTRATION	141
1.23	DRUG-DRUG INTERACTIONS	141
1.24	SPECIAL POPULATIONS.....	141
1.25	PEDIATRICS	141
1.26	ADVISORY COMMITTEE MEETING	142
1.27	LITERATURE REVIEW	142
1.28	POSTMARKETING RISK MANAGEMENT PLAN	142
1.29	OTHER RELEVANT MATERIALS	142
9.	OVERALL ASSESSMENT.....	142
1.30	CONCLUSIONS	142
1.31	RECOMMENDATION ON REGULATORY ACTION	143
1.32	RECOMMENDATION ON POSTMARKETING ACTIONS	143
1.32.1	Risk Management Activity	143
1.32.2	Required Phase 4 Commitments	143
1.32.3	Other Phase 4 Requests.....	143
1.33	LABELING REVIEW.....	143
10.	DESCRIPTION	143
1.33.2	OSAHS I	155
1.34	ALLERGIC REACTIONS.....	161

1.35	COMMENTS TO APPLICANT.....	181
o	EXAMINATIONS BEING CARRIED OUT BY THE PSYCHIATRY DIVISION AT THE FDA.....	182
o	REQUEST THE SUBMISSION OF ALL NARRATIONS (SERIOUS AND DISCONTINUATIONS) THAT DEALT WITH SKIN REACTIONS FROM ALL CLINICAL TRIALS DATING BACK TO THE ORIGINAL NDA SUBMISSION.....	182
11.	APPENDICES	183
1.36	APPENDIX A: REVIEW OF INDIVIDUAL STUDY REPORTS	184
1.36.1	Studies in OSAHS (3021 and 3025)	184
1.36.2	Study in Narcolepsy (3020)	194
1.36.3	Study in SWSD (3022)	200
1.37	APPENDIX B: COGNITIVE DRUG RESEARCH (CDR) SYSTEM.....	209
1.38	APPENDIX C: SHIFTS FROM BASELINE BASELINE TO ENDPOINT IN BLOOD CHEMSITRY TESTING FOR PATEINTS IN OSAHS PLACEBO-CONTROL PHASE 3 TRIALS.....	211
1.39	APPENDIX D: SELECTED BLOOD CHEMISTRIES CENTRAL TENDENCY ANALYSIS FOR PIVOTAL CONTROL TRIALS. 212	
1.40	APPENDIX E: SHIFTS FROM BASELINE TO ENDPOINT IN HEMATOLOGY INDICES FOR PATIENTS IN OSAHS PLACEBO-CONTROL PHASE 3 TRIALS	214
1.41	APPENDIX F: CHANGES IN EKG PARAMETERS (EXCLUDING QT) IN PLACEBO CONTROL PHASE 3 TRIALS. 215	
1.42	APPENDIX G: CHANGE IN STAGED SLEEP FROM BASELINE AT ENDPOINT ACROSS THE VARIOUS PLACEBO CONTROLLED PHASE 3 STUDIED DISORDERS.....	216
1.43	APPENDIX H CRITERIA FOR CLINICALLY SIGNIFICANT ABNORMAL VALUES IN SELECTED PERTINENT LABORATORIES.....	218

Appears This Way
On Original

EXECUTIVE SUMMARY

1.1 Recommendation on Regulatory Action

The application should be given an approvable action with a request for the following information:

- Because the Agency is currently evaluating submitted information on single possible case of Stevens Johnson Syndrome (SJS) in the clinical database for Provigil in a pediatric patient it is necessary to determine whether this is a case of SJS. If this is determined a prominent statement describing this event will need to be placed in product labeling for Nuvigil, despite the fact that Nuvigil will not be approved for use in the pediatric population. Final labeling, therefore, cannot be approved until our evaluation of that case is completed, and the larger issue of modafinil's **capacity to cause serious rash, is complete**. Moreover the Sponsor should search the entire clinical trials database using all appropriate search terms to identify any possible such cases, followed by a critical examination of all potential cases identified. The Sponsor should provide this division with all narrations (discontinuations and serious drug reactions) from the clinical trial database, dating back to the original NDA approval that contains any reference to skin adverse event.
- Because the Division has recently received a MedWatch report (Manufacturer #US016978), dated March 29, 2006, of a 31 year old man who died secondary to cardiac failure potentially resulting from multi-organ hypersensitivity syndrome related to modafinil treatment additional information on this potential adverse event should be requested. This would include more information on this case as well as an examination of both clinical trials and post-marketing database for other potential cases of multiorgan hypersensitivity.

Moreover an, integrated report of urinalysis data was not submitted till 4/6/06, which was too late in the review cycle to have been reviewed prior to this action. This will be reviewed in the next cycle.

The reader is referred the Clinical Pharmacology review for additional requested information.

Appears This Way
On Original

1.2 Recommendation on Postmarketing Actions

1.2.1 Risk Management Activity

Because of the potential confusion between Nuvigil and Norinyl (any strength) the division has requested that the Sponsor submit all medication error reports, actual or potential medication, as 15-day safety reports. The Sponsor has will be asked to include language in promotional materials describing the potential for this medication error.

Additional recommendation may be added following a complete examination of the issue of serious skin (and other allergic) reactions.

1.2.2 Required Phase 4 Commitments

Based upon PREA, studies in childhood narcolepsy will be required.

As per the pharm/tox review ~~_____~~ the carcinogenicity studies **b(4)** **b(5)** will be needed.

1.2.3 Other Phase 4 Requests

None.

1.3 Summary of Clinical Findings

1.3.1 Brief Overview of Clinical Program

Armodafinil is being proposed as a wakefulness promoting agent for patients with sleepiness associated with Obstructive Sleep Apnea Hypopnea Syndrome (OSAHS), Shift Work Sleep Disorder (SWSD) and narcolepsy. It is the pure R isomer of modafinil (Provigil), which is a racemate and is already approved for the same indication. The Sponsor is proposing a trade name of Nuvigil. This agent is being developed to be taken once daily, by mouth, to treat OSAHS and narcolepsy (150 mg and 250 mg) and once daily to treat SWSD (150 mg)

The complete phase 3 development program consisted of 4 pivotal (see table below) and two open label extension trials.

A summary of the features of the 4 pivotal studies can be found in the table below.

Disorder	Study number	Duration	Number of patients treated			
			Armodafinil (mg/day)			Placebo
			250	150	Total	Total
OSAS						
	Study 3021	12 weeks	131	131	262	392
	Study 3025	12 weeks	—	129	129	259
SWSD						
	Study 3022	12 weeks	—	123	123	245
Narcolepsy						
	Study 3020	12 weeks	67	64	131	194
Total			198	447	645	1090

The phase 3 open label extension trials (study 3023 and 3024), one of which (3024) served as an extension to controlled trials, studied 840 patients. A total of 245 healthy patients participated in earlier phase 1 and 2 pharmacokinetic and pharmacodynamic studies.

The total drug exposure at the time of the safety update is presented in the table below. This exposure is considered adequate, particularly when added to the previous clinical database for the racemate, Provigil.

Study drug exposure	≤100 mg (N=118)	150 mg (N=484)	200 mg (N=115)	250 mg (N=354)	All (N=1271)
Duration range, n (%)					
<2 weeks	12 (10)	15 (3)	3 (3)	5 (<1)	35 (3)
≥2 weeks and <1 month	22 (19)	25 (5)	5 (4)	35 (6)	87 (7)
≥1 month and <2 months	9 (8)	33 (7)	4 (3)	27 (5)	73 (6)
≥2 months and <3 months	5 (4)	90 (19)	3 (3)	36 (6)	134 (11)
≥3 months and <6 months	9 (8)	83 (17)	9 (8)	60 (11)	161 (13)
≥6 months and <9 months	10 (8)	28 (6)	17 (15)	56 (10)	111 (9)
≥9 months and <12 months	32 (27)	94 (19)	37 (32)	172 (31)	335 (26)
At least 12 months	19 (16)	116 (24)	37 (32)	163 (29)	335 (26)
Patient-years	62.92	283.05	89.64	410.30	845.91

1.3.2 Efficacy

The pivotal studies were double-blind, placebo-controlled, parallel arm, multi-institutional studies with an experimental phase of 3 months. Each study required statistical significance ($p < 0.05$) for both of two co-primary endpoints in order to conclude efficacy. One primary endpoint was an objective measure of sleepiness whereas the other was a physician scored subjective measure.

The objective measures in all studies consisted of an evaluation of sleep latency (measured in minutes). The endpoint sleep latency in narcolepsy and OSAHS studies were evaluated as the average of 4 test sessions (0900 to 1500) measured by the Maintenance of Wakefulness Test (MWT). The endpoint sleep latency in SWSD studies was evaluated as the average of 4 test sessions during a simulated evening work period (0200-0800) measured by the Multiple Sleep Latency Test (MSLT). The final evaluation utilized the difference between baseline and week 12 or final test session (ANCOVA/ANOVA).

The subjective co-primary endpoint measurement of sleepiness used in all studies was the Clinical Global Impression of Change (CGI-C) which **consists of the clinician's assessment of change in disease (sleepiness) severity relative to a baseline evaluation.** The CGI-C consists of the following categories and scoring assignments: 1=Very much improved; 2=Much improved; 3=minimally improved; 4=No change; 5=Minimally worse; 6=Much worse; and 7=Very much worse. The difference in the number of patients, who experienced at least a minimal improvement in condition at 12 week or final evaluation, was tested through chi-square test.

Secondary endpoints included other subjective measures of sleepiness including Epworth Sleepiness Scale (ESS) in SWSD and narcolepsy study and the Karolinska Sleepiness Scale (KSS) in the SWSD study. Sleep diaries were also maintained, the results of which were presented in a descriptive format.

Cognitive testing sub-scales in the Cognitive Drug Research System (CDR) were identified by **the Sponsor as "key secondary endpoints."** For 3 of the 4 studies (Studies 3020, 3021 and 3022 for narcolepsy, OSAHS and SWSD, respectively) the key secondary variable was the change from baseline to endpoint in the quality of episodic secondary memory. For one study (study 3025 on OSHAS), the key secondary endpoint was identified as the change from baseline to endpoint in the mean power of attention.

Examination of the CGI-C suggested a therapeutic effect. This is apparent from the table below that presents the percent difference in the number of patients who exhibited at least minimal improvement. Most of the differences between placebo and Provigil groups can be attributed to **patients with scores of "much" too "very much improved"**.

Disorder	NUVIGIL 150 mg*	NUVIGIL 250 mg*	Placebo
OSAHS I	71%	74%	37%
OSAHS II	71%	-	53%
Narcolepsy	69%	73%	33%
SWSD	79%	-----	59%

*Significantly different than placebo for all trials (p<0.01)

Examination of the latency to sleep evaluations indicated a statistically significant effect in each respective disorder at all doses studied. The table below summarizes such data.

Disorder	Measure	NUVIGIL 150 mg *		NUVIGIL 250 mg*		Placebo	
		Baseline	Change from Baseline	Baseline	Change from Baseline	Baseline	Change from Baseline
OSAHS I	MWT	21.5	1.7	23.3	2.2	23.2	-1.7
OSAHS II	MWT	23.7	2.3	-	-	23.3	-1.3
Narcolepsy	MWT	12.1	1.3	9.5	2.6	12.5	-1.9
SWSD	MSLT	2.3	3.1	-	-	2.4	.4

*Significantly different than placebo for all trials (p<0.01)

All final sleep latency evaluations were scored by a single central laboratory. Local technicians, however, were required to score the study in real time, so that they may awaken patients in studies once sleep criteria were achieved. Central scoring assigned maximal sleep latency to all sessions where sleep was not achieved including discordant sessions where the local, but not central, reader scored sleep. A large number of these discordant (or flawed) sessions were observed in studies using the MWT (OSAHS and narcolepsy). This occurred at a rate of 6.2% to 15.3% of the sessions in the various treatment groups. Incidence of these flawed sessions were equally distributed between drug and placebo. Because of this high rate the FDA requested a number of post-hoc analyses, including latency as determined by local readers. These post-hoc evaluations supported the efficacy conclusions for Nuvigil. This problem did not exist for the SWSD study where the MSLT was used an endpoint. This post-hoc analysis indicated that the problem of flawed sessions should not significantly effect the impetration of the data and therefore confirmed efficacy.

The conclusion of efficacy is further supported by secondary endpoint sleepiness rating scales, ESS and KSS which indicted statistically significant improvement. The conclusion of efficacy is also supported by the descriptive observation from **patient's diaries** that indicate reduced unintentional episodes of sleep and reduced **intentional napping**. **The information on "accidents and near misses," in these diaries did not always** demonstrate improvement with armodafinil treatment. Some studies did show some improvement in this measure in the armodafinil group.

Key secondary endpoints were identified to permit a labeling claim. The Sponsor was requested to present information that effects on these endpoints were independent of the wakfulness promoting effect of this agent. A satisfactory argument was not provided. **Nonetheless, "Power of Attention," a key secondary** endpoint for the OSAHS study 3025 was not statistically significantly improved. The second key secondary endpoint, Episodic Secondary Memory, was identified as a key secondary endpoint for studies 3020, 3021 and 3022 (narcolepsy, OSAHS and SWSD studies, respectively). This endpoint indicated a statistically significant effect, as compared to placebo, in two of these studies (3020 and 3022) but not for the OSAHS study 3021. Because of the lack of consistency of an effect and the lack of evidence that these

endpoints are not related to the improvement of wakefulness labeling for this effect is not recommended.

The magnitude of effects observed in the present study with Nuvigil are similar to effects observed with Provigil. There is no evidence that Nuvigil is superior to the already marketed Provigil.

1.3.3 Safety

No deaths were reported in the original NDA database. However, a single death was noted in the complete database in the safety update. This death occurred in an open label extension trial in a 59 year old male who was being treated for OSAHS. **The cause of death was “arterosclerotic cardiovascular disease.” Drug causality in this** is confounded by the fact that the patient was on a COX-2 inhibitor. The issue of potential cardiac ischemia was more completely evaluated in an examination of serious adverse events and discontinuations (see below).

There were 8 reported serious adverse events in the controlled database. The most common class that may potentially be linked to drug was that of **“psychiatric disorder.” Thus, 2 patients on drug** were noted to have serious psychiatric adverse events. . No such cases were reported for patients on placebo. One case was that of worsening of mood and personality disorder the other was a **case of worsening of depression with “suicidal intent” (walked** into the desert).

Another noteworthy serious adverse event in the control database was a single case of angioneurotic edema for a patient on drug. Another case of potential angioneurtic edema was observed in the open label database, and was not described as serious but as a reason for **discontinuation. While this case was described as “hypersensitivity, dysphagia and broncospasm” it likely represents a case of “angioneurtic edema.” There were no serious skin** reactions.

Serious adverse events in the original NDA open-label database included 2 cases of myocardial infarction. These occurred in patients with risk factors. Because of this, and the fact that it was not in the controlled database, it is difficult to attribute causality to drug. There were also 3 cases of chest pain in the uncontrolled database that do not contribute to causality as: 1) such events were not observed in the control database and included patients who possessed risk factors for cardiac disease, 2) chest pain may have been confounded with other cases in one case (e.g. gastric reflux), 3) definitive diagnostic testing (cardiac enzymes or EKG during the event) were not described that will allow cardiac attribution. This does not mean that in attribution is not possible but only that the data are not adequate to make such a contention.

Leading causes of discontinuations in the controlled database that may be attributed tot of armodifinil treatment, based upon of differences in incidence with placebo treated groups, included discontinuations for the following: 1) anxiety/ agitation/ nervousness/ irritability (drug 1.2% and placebo 0.3%), 2) depression (drug 0.6% and placebo 0.2%), 3) insomnia (drug 0.3% and placebo 0.0%), 4) headache/migraine (drug 2.2% and placebo 0.6%), 5) dizziness (drug

1.2% and placebo 0.0%), 6) nausea (drug 0.86% and placebo 0.22%, 7) cardiac flutter/palpitations (drug 0.78% and placebo 0.22%). Of note one case of depression (described above) was associated with suicidal intent and a latter case of suicide ideation was observed in the open label safety update. There were, however, no suicides. It is also noteworthy that there was insufficient evidence to link cardiac palpation/flutter to actual cardiac rhythm disturbance. This link, however, was not routinely investigated. Rash was of particular interest as a reason for discontinuation because of reports of potential serious skin reactions with Provigil. A similar percent of patients withdrew because of rash in placebo (0.45%) and drug (0.62%) in the controlled database. Some of these rashes did appear to be a potential allergic response to drug. Examination of narrations for discontinuations in the complete phase 3 database did not reveal any cases that appeared to represent a serious skin reactions (TEN, Stevens-Johnson or erythema multiforme).

Examinations of drug discontinuation resulting from potential cardiac ischemia (chest pain and EKG changes) did not reveal an obvious difference between placebo (0.67%) and drug (0.62%) treated groups in the control database.

Common adverse events table, derived from the Sponsors NDA, is presented in the table below. It includes any adverse event with a > 2% incidence in any dose group or placebo. Number (and percent) of patients are presented in this table.

Appears This Way
On Original

System organ class Preferred term	Number (%) of patients			
	Armodafinil		ALL	Placebo
	250 mg (N=190)	150 mg (N=147)	(N=445)	(N=145)
No. of patients with at least 1 AE	137 (69)	270 (60)	407 (65)	213 (48)
Cardiac disorders				
Palpitations	6 (3)	7 (2)	13 (2)	5 (1)
Gastrointestinal disorders				
Nausea	18 (9)	27 (6)	45 (7)	14 (3)
Diarrhea	7 (4)	19 (4)	26 (4)	8 (2)
Dry mouth	13 (7)	11 (2)	24 (4)	3 (<1)
Dyspepsia	2 (1)	14 (3)	16 (2)	2 (<1)
Constipation	4 (2)	4 (<1)	8 (1)	2 (<1)
General disorders and administration site conditions				
Chest pain	5 (3)	7 (2)	12 (2)	9 (2)
Fatigue	2 (1)	10 (2)	12 (2)	6 (1)
Thirst	5 (3)	1 (<1)	6 (<1)	1 (<1)
Pyrexia	4 (2)	0	4 (<1)	1 (<1)
Investigations				
Alanine aminotransferase increased	4 (2)	5 (1)	9 (1)	5 (1)
Aspartate aminotransferase increased	4 (2)	4 (<1)	8 (1)	4 (<1)
Metabolism and nutrition disorders				
Anorexia	6 (3)	3 (<1)	9 (1)	0
Decreased appetite	6 (3)	3 (<1)	9 (1)	0
Musculoskeletal and connective tissue disorders				
Back pain	4 (2)	11 (2)	15 (2)	8 (2)
Nervous system disorders				
Headache	46 (23)	63 (14)	109 (17)	39 (9)
Dizziness	10 (5)	20 (4)	30 (5)	8 (2)
Tremor	4 (2)	1 (<1)	5 (<1)	0
Psychiatric disorders				
Insomnia	12 (6)	18 (4)	30 (5)	5 (1)
Anxiety	10 (5)	18 (4)	28 (4)	4 (<1)
Depression	6 (3)	6 (1)	12 (2)	1 (<1)
Respiratory, thoracic and mediastinal disorders				
Cough	4 (2)	5 (1)	9 (1)	7 (2)
Pharyngolaryngeal pain	4 (2)	3 (<1)	7 (1)	3 (<1)
Skin and subcutaneous tissue disorders				
Rash	7 (4)	6 (1)	13 (2)	1 (<1)

Central tendency analysis of serum chemistry in the controlled database revealed slight average reduction in uric acid and small increase in GGT. Isolated significant elevations in GGT were observed slightly more frequently in the drug than the placebo group in the control database but these were not generally associated with increases in transaminase and never associated with bilirubin elevation. Sporadic significant elevations in ALT and AST were observed in the control database at a similar rate in drug and placebo groups but not associated with elevations in bilirubin. Three cases of significant elevations in bilirubin were observed in the complete phase 3 database, but these were not associated with transaminase elevations and patients tended to have elevated pre-drug bilirubin levels.

Central tendency analysis of blood counts did not reveal any significant average change between drug and placebo groups in the various indices of the CBC in the controlled phase 3 database. Significant outlier analysis indicated only a slightly more frequent incidence of significantly low WBC in the drug group of the controlled phase 3 database (0.8% drug versus 0.4% placebo). This was not expressed in significant outliers for absolute neutrophil counts. There were no serious hematological events reported although one patient was discontinued for reasons of "pancytopenia." **The WBC declined to 3.01, hematocrit to 31 and platelet to 108 from normal**

values. Some of these indices already showed a trend toward reduction over time before drug was started. Indices returned to normal following drug discontinuation.

Central tendency analysis for blood pressure in controlled trials showed only small average increases in mean systolic and diastolic blood pressure in patients receiving Nuvigil as compared to placebo (1.2 to 4.3 mmHg in the various experimental groups). There were also a slightly greater proportion of patients on Nuvigil requiring new or increased use of antihypertensive medications (2.9%) compared to patients on placebo (1.8%) in the phase 3 controlled database. Clinically significant outlier incidence between placebo and control groups was not very different. No cases of increased blood pressure were classified as a serious adverse event.

EKGs were performed at selected times at approximately T_{max} throughout the phase 3 studies. No significant difference between placebo and drug treated groups was observed in analysis of central tendency for QTcF and QTcB. No significant effects were observed by drug on PR interval and the QRS duration.

It is noteworthy, that there is no evidence that Nuvigil is better tolerated or safer than Provigil, which is marketed for the same indications.

Two recent significant observations have been made regarding modafinil, the racemate of armodafinil.

- A recent advisory committee (3/23/06) was convened regarding an NDA application for the approval of modafinil for the treatment of ADHD in the pediatric population. The principal reason this panel was convened was because of three suspicious rashes that may potentially represent a serious skin reaction. Only one of these rashes, as judged by the expert dermatologist, was considered suspicious for Stevens-Johnson syndrome. The **single episode of Steven's Johnson in the** pediatric clinical study database for the modafinil NDA database, however, would make the incidence in the pediatric population markedly high.
- A recent Provigil postmarketing report was of some concern (Manufacture # US016978). Thus, a death was observed in a 31 year old male that appeared to result from a multiorgan hypersensitivity reaction (pathologically proven). This case was first observed 3 days after Provigil treatment was started and expressed itself with rash, cardiac symptoms and eosinophilia. The patient was also on Trileptal and Zoloft for years.

1.3.4 Dosing Regimen and Administration

The recommended dose of Nuvigil is 150 or 250 mg given once a day for patients with OSAHS or narcolepsy, and 150 mg/day for patients with SWSD. These doses are based upon those investigated in the controlled trials. While the doses appear to produce an acceptable therapeutic benefit Vs risk a full dose response investigation was not performed in controlled studies.

Clinical Review
Norman Hershkowitz, MD, PhD
21,875 (000)
Nuvigil (armodafinil)

1.3.5 Drug-Drug Interactions

The reader is referred to the Clinical Pharmacology Review.

Appears This Way
On Original

INTRODUCTION AND BACKGROUND

1.4 Product Information

Armodafinil is being proposed as a wakefulness promoting agent. It is the pure R isomer of modafinil (Provigil), which is the racemate and already approved for the same indication. The Sponsor is proposing a trade name of Nuvigil.

In animal studies armodafinil appears to have a similar wakefulness promoting effect as observed for modafinil. The Sponsor contends that modafinil is distinct from the more classical stimulants, such as amphetamine, in that, although it inhibits dopamine uptake like these stimulants, modafinil does not directly promote dopamine release.

The Sponsor is proposing that it be approved for the treatment of the following disorders at the indicated regimens:

- Obstructive Sleep Apnea Hypopnea Syndrome (OSAHS): dose of 150 or 250 mg qD.
- Narcolepsy: 150 or 250 mg qD.
- Shift Work Sleep Disorder: 150 mg qD

1.5 Currently Available Treatment for Indications

Other than Provigil only methylphenidate hydrochloride and dextroamphetamine are approved for use in the United State for the treatment of excessive sleepiness associated with narcolepsy.

Other drugs are used by the medical community, off label, for the treatment of excessive sleep associated with narcolepsy include pemoline, and methamphetamine.

All agents that have been approved or used off label for the treatment of narcolepsy except Provigil, are schedule II. Provigil is schedule IV.

At present only Provigil is indicated and labeled for the treatment of OSAHS and narcolepsy.

1.6 Availability of Proposed Active Ingredient in the United States

As noted previously, armodafinil is the pure R isomer of the presently available racemate, Provigil.

1.7 Important Issues With Pharmacologically Related Products

A recent advisory committee (3/23/06) was convened regarding an application for the approval of modafinil for the treatment of ADHD in the pediatric population. The principal reason this panel was convened was because of three suspicious rashes that may potentially represent a serious skin reaction. Only one of these rashes, as judged by the expert dermatologist, was considered suspicious for Stevens-Johnson syndrome. This case is described below:

A 7 year old male was noted to have a sore throat, fever and mild rash by day 16 of treatment with modafinil 340 mg/day. Amoxicillin was started on day 17, but treatment was limited to a single dose. The modafinil stopped by day 17. By day 19, the rash was spreading, and continued to progress, with blistering, peeling, and mucosal involvement (lips and urethral meatus). A dermatologist made a diagnosis of Stevens-Johnson. The rash appeared to resolve by day 30.

No apparent previous serious skin reactions consistent with Steven-Johnson were reported for Provigil in prior adult clinical trials reported by the Sponsor. Postmarketing surveillance data requested before the advisory meeting revealed **5 cases of Steven's Johnson syndrome/erythema multiforme** resulting in hospitalization. Examination of use and background data led Dr. Racoosin, DNP safety team leader, to conclude in a memo (4/6/06) that **"the reporting risk for SJS cases [in adults] with modafinil over the period 2002-2005 approaches but does not meet the lower limit of the background rate for SJS."**

The single episode of Steven's Johnson in the pediatric clinical study database for the NDA, however, would make the incidence in the pediatric population markedly higher than the background rate. Dr. Racoosin concluded in her memo that **"the appearance of serious skin rash in the pediatric ADHD development program is of significant concern, and further study in this population would be prudent."**

The medical reviewer for the NDA, at the advisory meeting, suggested that the pediatric/adult difference in incidence may result from the increased exposures experienced in the ADHD studies observed for modafinil in general, and the sulfone metabolite specifically. These were substantially higher than that observed in at recommended adult therapeutic dosages because of differences in dosage and metabolism. Some at the advisory meeting raised doubt that exposure could be a factor in idiosyncratic skin reactions. Because of the single pediatric serious skin reaction in such a small database the advisory panel recommended 8 to 1 not to approve modafinil for ADHD in children until more data on risk were obtained. It was recommended that the Sponsors accrue experience on an additional 3,000 pediatric patients.

These skin reactions have been a subject of a CBE labeling supplement. The division presently intends to meet to consider labeling changes for Provigil based upon these data. The division should request complete narratives for all discontinuations (an serious as well) narratives to confirm that none of these may have been potential cases of serious skin reactions that may have not been labeled as such.

In addition to the above a recent Provigil postmarketing report was of some concern (Manufacture # US016978). Thus, a death was observed in a 31 year old male that appeared to

result form a multiorgan hypersensitivity reaction (pathologically proven). This case was first observed 3 days after Provigil treatment was started and expressed itself with rash, cardiac symptoms and eosinophilia. The patient was presumably on Trileptal and Zoloft for years. The division should request more information to help determine causality in this case. This information should include, but not limited, information on when Provigil was discontinued and confirmation that the patient was on Trileptal, **which is associated with this disorder, for “many years.”** **The Sponsor should examine their complete** database (clinical trials and postmarketing) to determine if there are any similar cases. The examination should be exhaustive and will require a careful examination of a number of clinical signs and symptoms as this disorder may effect a number of organ systems.

1.8 Presubmission Regulatory Activity

The division has meet with the Sponsor in both end of phase 2 and p[re NDA meetings and had approved the general research program,

Other Relevant Background Information

SIGNIFICANT FINDINGS FROM OTHER REVIEW DISCIPLINES

1.9 CMC (and Product Microbiology, if Applicable)

No major issues were raised, but the reader is referred to the CMC review.

1.10 Animal Pharmacology/Toxicology

This division’s pharm/tox reviewer have found inadequate inadequacies in the carcinogenicity studies that will require an approvable action. The reader is referred to the pharm/tox review.

DATA SOURCES, REVIEW STRATEGY, AND DATA INTEGRITY

1.11 Sources of Clinical Data

The data base includes 1169 unique patients with one of the three studied sleep disorders who have received armodafinil in a total of 6 phase 3 studies. Data cut-off for the presentation in the initial NDA submission was 12/15/04. Of the 6 phase 3 studies, 4 were pivotal double-blind placebo-control studies. Two of the placebo-control studies were performed in patients with OSAHS and one in patients with narcolepsy and another in patients with SWSD. The table below presents summary information on the double-blind placebo-controlled studies.

Disorder	Study number	Duration	Number of patients treated				
			Armodafinil (mg/day)			Placebo	Total
			250	150	Total		
OSAHS							
	Study 3021	12 weeks	131	131	262	392	
	Study 3025	12 weeks	—	129	129	259	
SWSD							
	Study 3022	12 weeks	—	123	123	245	
Narcolepsy							
	Study 3020	12 weeks	67	64	131	194	
Total			198	447	645	1090	

The remainder of the phase 3 studies included 2 ongoing long term open label studies. These are summarized in the table below. Dosages used in these studies were 100 to 250 mg/day. As apparent patients in these studies suffered any of the three sleep disorder. Patients in study 3024 were required to have participated in previous phase 3 double-blind placebo controlled studies. Patients in 3023 were not derived from prior double-blind studies.

Study number	Number of patients treated by sleep disorder			Total
	OSAHS	SWSD	Narcolepsy	
Study 3023 ^a	164	107	48	319
Study 3024 ^b	407	42	72	521
Total	571	149	120	840

An additional 245 healthy subjects received armodafinil in a variety of PK, bioequivalence and PD studies. These are briefly summarized in the table below.

Study number	Study design	Duration of treatment	Armodafinil dose	N ^a
Study 1023	Open-label, 2-way crossover, bioequivalence	Single dose (x2)	250 mg	30
Study 101	Double-blind, placebo-controlled, pharmacokinetics	Single dose ^c	50–400 mg	30
Study 102	Double-blind, placebo-controlled, pharmacokinetics	14 days	50–400 mg	37
Study 103 ^b	Double-blind, placebo- and active-controlled, pharmacodynamics	Single dose	100–300 mg	71
Study 1021	Open-label, 2-way crossover, drug interaction	Single dose	400 mg	24
Study 1022	Open-label, drug interaction	31 days	250 mg	24
Study 1025	Open-label, drug interaction	29 days	250 mg	29
Total				245

1.12 Tables of Clinical Studies

See previous section.

1.13 Review Strategy

All studies were used for a determination of safety. Only the 4 pivotal double-blind placebo-control studies were used for a determination of efficacy.

1.14 Data Quality and Integrity

Five study sites were inspected by DSI. DSI concluded:

There were no significant deficiencies found with the conduct of the above 5 studies by the principal investigators. All studies appeared to have been adequately conducted with a few protocol deviations noted at the Lahmeyer site. There was no evidence of under-reporting of adverse events at any site. However, because of the protocol design, much of the study data could not be verified by inspection. Of primary significance is that all source PSG [sleep latency measurements through MSLT and MWT] readings obtained at the study sites were sent to a third party for initial analysis, and the analysis results were subsequently sent only to the sponsor for computation of the primary efficacy endpoint; mean sleep latency from the MWT values. Study sites were never given the results of the third party analysis of the PSG data, or the results of the computation of the mean sleep latency from the MWT values, so it was not available on site for verification during the inspection.

Because of irregularities identified in the sleep latency dataset, and described in the section on endpoints, DSI was requested to investigate the central reading site:

No obvious violations were observed.

b(4)

1.15 Compliance with Good Clinical Practices

There were a number of minor violations observed by DSI (see DSI report) and by the Sponsor, (see protocol description in Appendix A). Collectively these were not of adequate magnitude to impact upon the conclusions made in this NDA review.

1.16 Financial Disclosures

Financial disclosure forms were filed. A few of the investigators disclosed holding (stocks) in Cephalon and significant consulting fees received from Cephalon. The argument is made that these would unlikely effect the results as they constitute a small percent of the total number of investigators participating in these large multiinstitution trials.

CLINICAL PHARMACOLOGY

Clinical Pharmacology reviewer found the present application acceptable. Noted is adequate bridging studies that were presented to allow the use of Provigil (racemic modafinil) in the labeling.

INTEGRATED REVIEW OF EFFICACY

1.17 Indication

The Sponsor is seeking approval of **Nuvigil in the indication “to improve wakefulness in patients with excessive sleepiness associated with obstructive sleep apnea/hypopnea syndrome, narcolepsy and shift work sleep disorder.”**

1.17.1 Methods

A total of 4 pivotal double-blind, placebo-control, multiinstitutional studies were performed. Two examine efficacy in OSAHS (studies 3021 and 3025) and 1 in each of the two remaining disorders, SWSD (study 3022) and narcolepsy (study 3020). The table below presents a summary of the numbers of patients and doses studied for these studies.

Sleep disorder	Number of efficacy-evaluable patients (full analysis set) ^a			
	Armodafinil			Placebo
	Study number	250 mg/day	150 mg/day	
OSAHS				
Study 3021	121	120	241	124
Study 3025	—	116	116	120
SWSD				
Study 3022	—	112	112	104
Narcolepsy				
Study 3020	60	58	118	58

^aThe full analysis set includes patients who received at least 1 dose of study drug, had a baseline and at least 1 postbaseline Maintenance of Wakefulness Test assessment (narcolepsy and OSAHS) or Multiple Sleep Latency Test assessment (SWSD), and at least 1 postbaseline Clinical Global Impression of Change assessment.
 OSAHS=obstructive sleep apnea/hypopnea syndrome; SWSD=sleep work shift disorder.

1.17.2 General Discussion of Endpoints

1.17.2.1 Primary Endpoint

1.17.2.1.1 Sleep Latency Evaluation

Each study contained two co-primary endpoints. One endpoint was an objective measure of sleepiness whereas the other was a physician scored subjective measure. The Sponsors was required to win on both measures.

The objective measures in all studies consisted of an evaluation of sleep latency. This was measured by a daytime Maintenance of Wakefulness Test (MWT) for the narcolepsy and OSAHS studies and by a Multiple Sleep Latency Test (MSLT) for the SWSD study.

The MSLT, used in the SWSD study, uses EEG to determine sleep latency and is generally **considered the "gold standard" for the measurement** of sleepiness. The test is designed by placing the patient in a bed in a darkened room with the instructions to sleep. **Four MSLT sleep latency tests were performed during a "simulated work period" and averaged for a final value.** These evaluations occurred the night following the last night of the shift work for that week. The MSLT was performed at 0200, 0400, 0600 and 0800 during these study sessions. An additional earlier assessments was made at 2,400 but was used in secondary endpoint analysis (see below). Latency is defined as the time of onset of 3 consecutive epochs of stage 1 sleep or a time to a single epoch of stage 2, 3, 4 or REM sleep. Epochs were scored using the 50% (or 16-second)

rule. All MWTs were scored by an independent evaluator at the _____
_____. Sleep latency was measured as the elapsed time from lights-out to the first epoch scored as sleep. If patients fall asleep they were immediately woken and maintained awake. The MSLT change from baseline to the week 12, or last visit, was used as the primary endpoint. Latencies of 9 minutes and greater is generally considered normal, Latencies of 5-8 minutes is considered borderline and latencies of less than 5 minutes is considered pathological.

b(4)

The MWT, used in studies on OSAHS and narcolepsy, uses EEG to determine sleep latency onset during a series of test sessions at scheduled times throughout the day (0900, 1100, 1300 and 1500). Note two latter additional MWTs assessments (1700 and 1900) were made but only used as part of the secondary endpoint analysis (see below). This test is similar was the MSLT except patients are instructed to remain awake in a darkened room while in a semi-reclined position during a test session lasting 20 minutes for narcolepsy and 30 minutes for OSAHS. For the primary endpoint evaluation the MWT was performed at baseline and following 12-weeks of treatment during the double-blind phase. Sleep latency for the 4 sessions were then averaged to obtain a final value. If a patient did not sleep within 20/30 minutes, that particular session was considered complete and a maximal value of 20/30 minute value was entered as the latency. Scoring was similar to that used for the MSLT. If patients fall asleep they were immediately woken and maintained awake. Sleep latency was measured for the 4 sessions and averaged for the final sleep latency. Some investigators believe that the MWT is more sensitive to certain types of therapeutic manipulations and may better represent the risks of “unintentionally falling asleep when trying to stay awake. Normative values have not been as clearly defined as that for the MSLT but one author has described normal values based upon values greater than 2 standard deviations below the mean in a normal population¹. Thus for a 20 minute sleep trial the cut-off for being to sleepy is 10.9 minutes. This is slightly longer than that for the MSLT. For a 40 minute sleep trial it is 12.9 minutes. MWT and MSLT do not always correlate, though they are considered to measure some aspect of sleepiness. The MWT, however, has been shown to respond to therapies and is considered by some to have the advantage face validity.²

1.17.2.1.1.1 Post submission Problems with Sleep Latency Dataset

In the process of review of the application the statistical reviewer noted, in examining the data sets, that a large number of sleep test sessions were prematurely terminated and assigned a value of **20 minutes in the final central analysis: these will be referred to as “flawed sessions.”** It was discovered that the reason for this was that local on-site sleep technicians were instructed to awaken patients and terminate the session once they achieved the criteria of sleep (described above). This was done to prevent the interference of “**unnecessary sleep**” with subsequent readings of latency. This problem was not noted in the study reports. Because of this the division requested clarification of a number of issues and additional post hoc analysis to

1 Doghramji, K. et. al. EEG Clin. Neurophys., 103: 554- (1997).
2 Poecta, J.S et al Chest 101: 893-897 (1992)/

determine how this problem might affect the final study outcome. The response to this initial inquiry was received by this division on 10/31/05.

Questions to the Sponsor revealed the following:

- The studies were performed at established sleep study centers with trained technician. It was the investigators responsibility to ensure the technicians were **well trained. Technician's scored** the study, in real time, based upon the Rechtschaffen and Kale rules. These are universal criteria for such studies. There was no requirement for over-read by a sleep specialist. Monitors would visit to assure compliance. The criteria used are described in the protocol and were **presented to technicians at a "Global sleep technicians meeting."** DVDs of these meeting were also made available to centers. This would suggest that data based upon the technicians read should carry some weight in a decision of therapeutic decision making (see below). There are, however, some degree of uncertainty with regard to the actual time to be used based upon the local reading (see the third bulleted item).
- The Sponsor was asked to explain the discrepancy between local and central reading that led to this problem. The Sponsor noted that sleep scoring is well **known to be "subject to individual interpretation,** and it is well recognized that a sleep stage identified for a given epoch may differ between scorers. In addition, periods of rapid sleep stage transition, such as wake to stage-1 sleep, and the presence of fragmented sleep as seen in patients with sleep disorders may lead to **greater variability between scorers."** Muscle artifacts may additionally contribute to variability. The Sponsor refers to a published study that describes a 75% inter-laboratory concurrence³. The Sponsor feels that the present study is consistent with this, noting a 6.7 to 12.5% incidence of truncated naps in the MWT and 1.1% in the MSLT. Examination of the presented manuscripts by this reviewer suggests that the differences between local and central readings were not far from that which were expected from these published reports.
- It was assumed that the local termination were made because of the achievement of sleep criteria. The Sponsor was asked whether there was any documentation that indeed the sessions were terminated because of this. To this the Sponsor noted that the only notes maintained were **"technical notes."** **These included the following information:** "pages of calibration, first and end epoch of each nap (page number and time), lights out (24-hour clock time), information that contributes to interpretation, environmental noise, comments by patients, and **changes in calibration."** Notes, however, "were not to indicate anything that might

³ Danker-Hopfe H, Kunz D, Gruber G, Klösch G, Lorenzo JL, Himanen SL, Kepm B, Penzel T, Röschke J, Dorn H, Schlögl A, Trenker E, Dorffner G. Interrater reliability between scorers from eight European sleep laboratories in subjects with different sleep disorders. *J Sleep Res* 2004;13(1):63-9; Norman RG, Pal I, Stewart C, Walsleben JA, Rapoport DM. Interobserver agreement among sleep scorers from different centers in a large dataset. *Sleep* 2000;23(7):901-8.

bias the centralized reader (e.g., snore, sleep latency value, warnings, epoch score).” **This meant that there is no** documentation of scoring by the local technicians. But, the Sponsor notes that technicians went through training on when to terminate the sleep. The design is understandable had the original central reading succeeded. It, however, complicates the use local readers termination period. That is, it leaves some uncertainty that termination was performed because of the achievement of sleep criteria.

To determine the scope of the present problem the Sponsor was asked to provide information on the number of flawed sessions. Below is a table (from the FDA’s statistical review) listing flawed sessions in all phase 3 double blind studies. It is most obvious that the incidence of flawed sessions in study 3022 (SWS study) is remarkably lower. This is likely the result of the fact that this study utilized the MSLT whereas other studies used the MWT. MSLT EEG data are less likely to be contaminated by EMG activity as patients are trying to sleep in a dark room rather than remain awake in a lighted room. EMG may obfuscate EEG pattern and result in more ambiguous scoring resulting with consequent discordant scoring between independent readers. The remainder of flawed sessions in critical sessions (first 4 sessions at baseline and endpoint treatment testing) varied between 6.2% to 14.9%. This, as noted above, is within the range of concurrence between independent readers. The incidence of flawed sessions appears somewhat lower in study 3022 (OSAHS) than the other two remaining studies (OSAHS and narcolepsy). Flawed sessions incidence within studies appears to be relatively randomly distributed across treatment arms.

summary	Protocol	CEP-10953 250 MG	CEP-10953 150 MG	Placebo	Total Overall
Pooled counts in all visits	3020	7.9% (109/1388)	15.3% (203/1328)	14.7% (195/1330)	12.5% (507/4046)
	3021	7.3% (197/2706)	6.2% (168/2696)	6.6% (190/2880)	6.7% (555/8282)
	3025		10% (268/2676)	13.5% (367/2723)	11.8% (635/5399)
	3022		1.5% (30/2060)	0.6% (12/1930)	1.1% (42/3990)
pooled counts in critical visits	3020	9.2% (66/715)	14.1% (98/694)	14.9% (103/692)	12.7% (267/2101)
	3021	8.6% (124/1438)	6.8% (98/1437)	6.2% (92/1481)	7.2% (314/4356)
	3025		9.6% (133/1382)	14% (199/1422)	11.8% (332/2804)
	3022		0.8% (9/1120)	0.2% (2/1045)	0.5% (11/2165)

The table below presents percent of patients affected by at least one flawed session in the first three naps of testing days. Again it is apparent that effected patients are relatively evenly distributed across treatment arms. But, also of note is the fact that in most cases a similar percent of patient’s scores were affected at baseline and endpoint measures.

Appears This Way
 On Original

Protocol	Visit	CEP-10953 250 MG	CEP-10953 180 MG	Placebo	Total Overall
3020	Baseline	16.7% (10/60)	34.5% (20/58)	31% (18/58)	27.3% (48/176)
	Endpoint	28.3% (17/60)	34.5% (20/58)	41.4% (24/58)	34.7% (81/178)
3021	Baseline	18.2% (22/121)	22.5% (27/120)	18.5% (23/124)	19.7% (72/365)
	Endpoint	21.5% (26/121)	17.5% (21/120)	12.1% (15/124)	17% (62/365)
3025	Baseline		21.6% (25/116)	31.7% (38/120)	28.7% (63/238)
	Endpoint		22.4% (26/116)	30% (38/120)	29.3% (62/238)
3022	Baseline		0% (0/112)	0% (0/104)	0% (0/216)
	Endpoint		4.5% (5/112)	1% (1/104)	2.8% (6/216)

In addition to questions the division asked for a number of post-hoc analyses (see analyses 1 to 6 below) in an attempt to correct for the problem of flawed sessions. These were received in the submission of 10/31/05. However additional "worst case" analyses were requested by statistics and received on 12/16/05 (see analyses 7 and 8, below). These are described and enumerated below. The results in this report will refer back to this enumeration.

- **Analysis 1.** Use the latency to sleep as determined by local site: i.e. based upon time of awakening at local site.
- **Analysis 2.** Use the last observation carried forward (LOCF) for endpoint visits with a flawed nap session. Thus, if there is a flawed score in any one of the four naps at the endpoint visit, use the next earliest visit where all four scores are unflawed (e.g., use week 8 if week 12 is incomplete because of a single flawed score or week four if week 12 and 8 are incomplete). If all experimental period naps are incomplete (weeks 4, 8, and 12), use the unflawed experimental median endpoint score for the endpoint visit for the combined placebo/drug groups to replace that **patient's endpoint value**. If one of the four baseline naps is flawed, use the median combined placebo/drug baseline value of unflawed data to replace **that patient's baseline value**.
- **Analysis 3.** In a visit with a flawed nap session, replace each flawed nap score with the next later unflawed nap score to obtain an average of 4 nap sessions (i.e. if one of the four naps is flawed, replace with the fifth and then sixth if there is a problem with the fifth). If two naps are flawed, replace with the next 2 unflawed naps. If there remains no unflawed nap to serve as a replacement, use the unflawed median of combined drug/placebo for that single nap score replacement (baseline median for baseline missing scores and endpoint experimental median for endpoint missing scores).
- **Analysis 4.** Discard flawed nap sessions that occur in the first 4 naps and replace with the median of unflawed scores for the combined placebo/drug (first 4 naps) for that particular experimental stage. Baseline values should be replaced with baseline medians and experimental endpoint with experimental endpoint median.
- **Analysis 5.** Discard data and do not analyze patients with any of the first 4 nap sessions flawed at baseline or final endpoint measure.
- **Analysis 6.** Average only nap sessions that are unflawed in the first four naps.
- **Analysis 7. Worst case analysis**
 - **Analysis 7.1.** In placebo group, replace each flawed session at baseline with the lowest sleep latency of six baseline sessions of each corresponding patient (take consideration of both flawed and unflawed sessions but count actual awakening time for sleep latency if it is a flawed session) and replace each flawed session at endpoint with highest sleep latency of unflawed sessions from six endpoint sessions of the corresponding patient. In treatment groups, replace each flawed session at baseline with the highest sleep latency of unflawed sessions from six baseline session of

- the corresponding patient and replace each flawed session at endpoint with lowest sleep latency of six endpoint sessions (take consideration of both flawed and unflawed sessions but count actual awakening time for sleep latency if it is a flawed session). This analysis biases toward an increase in latency in the placebo group and decrease latency in the armodafinil group following treatment.
- **Analysis 7.2.** Repeat analysis 7.1), but instead of using all six sessions, use the first 4 sessions.
 - **Analysis 8.** Most conservative worst case analysis
 - **Analysis 8.1.** In placebo group, replace each flawed session at baseline with the lowest sleep latency of six baseline sessions of the corresponding patient (take consideration of both flawed and unflawed sessions but count actual awakening time for sleep latency if it is a flawed session) and replace each flawed session at endpoint with full length of study session (20 or 30 minutes). In treatment group(s), replace each flawed session at baseline with full session length (20 or 30 minutes) and replace each flawed session at endpoint with lowest sleep latency of six endpoint sessions (take consideration of both flawed and unflawed sessions but count actual awakening time for sleep latency if it is a flawed session). This analysis biases even more than analysis 7 as it chooses the maximal awake time for placebo treatment period and armodafinil baseline treatment.
 - **Analysis 8.2.** Repeat analysis 8.1), but instead of using all six sessions, use the first 4 sessions.

Note that throughout this review the original analysis and post-hoc analyses are presented and discussed for the primary endpoint. Secondary endpoint sleep latency evaluations utilize solely the original NDA evaluation.

1.17.2.1.2 Clinical Global Impression of Change (CGI-C)

The subjective co-primary endpoint measurement of sleepiness used in all studies was the Clinical Global Impression of Change (CGI-C) which consists of the clinician's assessment of change in disease (sleepiness) severity relative to a baseline evaluation. Baseline severity of disease was assessed using the Clinical Global Impression of Severity of Illness (CGI-S), which consists of the following categories and scoring assignments: 1=Normal (shows no signs of illness); 2=Borderline ill; 3=Mildly (slightly) ill; 4=Moderately ill; 5=Markedly ill; 6=Severely ill; and 7=Among the most extremely ill patients. For the CGI-C the clinician assessed the change from this baseline. The CGI-C consists of the following categories and scoring assignments: 1=Very much improved; 2=Much improved; 3=minimally improved; 4=No change; 5=Minimally worse; 6=Much worse; and 7=Very much worse.

1.17.2.2 Secondary Endpoints

Although generally the same secondary endpoints were used amongst the different studies there were small differences. The table below presents a summary as to the different tests used in the various studies. All endpoints were measured at 4, 8 and 12 weeks and, except were used as a primary or key secondary endpoint, at the final endpoint analysis (week 12 or last observations).

	OSAHS	SWSD	Narcolepsy
--	-------	------	------------

Clinical Review
 Norman Hershkowitz, MD, PhD
 21,875 (000)
 Nuvigil (armodafinil)

	Study 3021	Study 3025	Study 3022	Study 3020
ESS	X	X	X	X
Epworth Sleepiness Scale (ESS)	X	X		X
Karolinska Sleepiness Scale (KSS)			X	
Brief Fatigue Inventory (BFI)	X	X	X	X
Cognitive Drug Research System (CDR)	X	X	X	X
Power of attention from CDR at final endpoint	X	X*	X	X
Episodic memory from CDR at final endpoint	X*	X	X*	X*
Average of the early first 4 MWT assessments.	X	X		X
Average of last 3 MWT assessments (1500, 1700, and 1900)	X	X		X
Average of last 4 MSLT assessments (0200 to 0800)			X	
Profile of all 5 MSLT assessments (2400-0800)			X	
CGI-C	X	X	X	X
Change in daytime sleepiness based upon diary.	X	X	X	X (nighttime working sleepiness)

*Key Secondary endpoints: although this endpoints was analyzed at all time points (4, 6, 8 12 weeks and final endpoint), only the final endpoint measure was used when this was used as a key secondary endpoint.

A single key secondary endpoints derived from different aspects of the Cognitive Drug Research System (CDR) was specified for each of the 4 studies.

At a meeting with the Sponsor in 10/22/04 the Sponsor queried whether the division would

~~_____~~ The Division responded by noting that ~~_____~~ Cephalon would need to: "1) demonstrate that this endpoint represents a domain measurement that is independent of the primary endpoint, 2) justify the validity and reliability of this endpoint, 3) present a reasonable statistical plan for its evaluation (e.g., including any correction for multiple comparisons)." The Sponsor was told that at first blush it appears to this Division that a therapeutic agent that improves wakefulness should also increase attention, i.e. wakefulness and attention are not independent domains. Cephalon was advised to submit an argument justifying the endpoint (points 1 and 2) before the NDA submission or at the time of submission and a statistical analysis plan prior to data is unblinding.

b(4)

The CDR was performed in all patients; however, sub-scales of the CDR were identified as a key secondary endpoint in different studies. The Sponsor noted in the present submission in support of validity that the “**the CDR system has been** used in over 600 clinical studies worldwide and has been the subject of over 140 papers, chapters, **published abstracts, and presentations.**” The full CDR and its subscales are more thoroughly described in Appendix B. According to the Sponsor presents references that support their contention that the CDR has been extensively validated and shown to be reliable, and measures psychological constructs that it was designed to measure. The Sponsor describes two studies, one in 272 healthy patients and one in a population of Lewy body dementia to support this contention. Some references examine pharmacological agents effect on the CDR . Some of these demonstrated anticholinesterase improvement in CDR score in dementias. While references are provided by the Sponsor there is not a thorough **narrative that proves the Sponsor’s** contention regarding validity and reliability. A cursory exam of the publications would indicate that this may be the case, but a through evaluation of the power of the data in the publications is beyond the scope of this review. Nonetheless, the Sponsor fails to present evidence that an effect on the CDR would not a result from increased wakefulness.

The two key secondary endpoints derived from the Cognitive Drug Research are described as follows:

- **Episodic Secondary Memory derived from Cognitive Drug Research System CDR System:** For 3 of the 4 studies (Studies 3020, 3021 and 3022 for narcolepsy, OSAHS and SWSD, respectively) the key secondary variable was the change from baseline to endpoint in the quality of episodic secondary memory derived from the tests of memory from the CDR system. Quality of episodic secondary memory tasks included immediate (number of words recalled correctly) and delayed word recall, word recognition, and picture recognition (the number of items correctly recognized and the number of items not previously presented that were correctly rejected). According to literature cited by the Sponsor this scale measures the ability to store, hold and retrieve information of an episodic nature (i.e. an event, a name, an object, a scene, an appointment).
- **Power of attention from CDR:** The key secondary endpoint in one study (study 3025 on OSHAS) was the change from baseline to endpoint in the mean power of attention (computed as the speed from the simple reaction time test plus the speed from the digit vigilance task plus the speed from the choice reaction time test) from the tests of attention from the CDR system. According to the Sponsor **the power of attention “reflects the intensity of concentration at a particular moment: the faster the response, the more processes that are being brought to bear upon the test.”**

Other Secondary endpoints are noted below:

- **The Epworth Sleepiness Scale (ESS):** The ESS is a subjective patient evaluation questioner of sleepiness. This evaluation asks patients for their propensity to fall asleep (on a scale from 0 to 3) in 8 common daily situations (e.g., sitting and reading, talking to

someone, while stopped in traffic in a car, watching television). The total potential score is 24. Lower scores indicate a lower degree of sleepiness. A score of 11 to 16 is considered a moderate degree of sleepiness and greater than 16 is considered excessive sleepiness.

- **Karolinska Sleep Scale (KSS):** The KSS is a patient rated scale that measures sleepiness on a scale of 1 to 9 with 1 being **“very alert”** and 9 being **“very sleepy.”** A score of 7 or greater is considered pathological. It has been validated by electrophysiological sleep parameter studies. It was conducted 8 times throughout the simulated night shifts before each MSLT and PVT. The average of these scores was used for comparison.
- **Brief Fatigue Inventory (BFI).** This is a subjective questioner containing 9 items that are rated numerically from 0 to 10, with 10 **indicating the “as bad you can imagine”** and 0 **indicating “none”**. **Questions include degree of fatigue “now, usual degree over the last 24 hours, worse fatigue over 24 hours and how fatigue interfered with various activities over 24 hours.** The scale is based upon a scale used in pain (BPI). Reliability and validity has been demonstrated in studies comparing normal controls to cancer patients. A mean score from 0 to 10 is derived. Scores may be interpreted as follows: mild 1-3, moderate 4-6 and severe 7-10 fatigue. This reviewer could not find any studies examining the validity or reliability of the BFI in dysomnias.
- **Patient Diary:** Patient diaries allowed the recording of unintended sleep episodes and naps, the number of caffeinated, the number of mistakes, accidents, or near misses and effect on nighttime sleep.

1.17.3 Study Design

The 4 pivotal studies were of a placebo-control, double-blind design and very similar design to used in the approval of the racemate Provigil. This reviewer believes that this study is of adequate design to detect a therapeutic effect.

The study was of adequate duration, 3 months. Moreover, secondary endpoint evaluations (serial evaluations during the experimental phase) allowed for the examination of potential tolerance to this drug's therapeutic effect.

The dose selection (150 and 250 mg/day) was based upon two pharmacokinetic studies: study 101 and 102 which examined single and multiple doses. The dose selection was also based upon a single pharmacokinetic/pharmacodynamic/tolerance study (study 103). Thus, study 102 examined treatment with 50 to 400mg of armodafinil daily over a period of 14 days. On the basis of this study the dose of 300 mg/day was considered the maximum tolerated dose. Study

103 was a double blind study that examined single placebo dose versus single armodafinil (100, 150, 200 and 300 mg/day) and Provigil (200 mg/day) dose in young normal men (n=108) undergoing sleep deprivation. The study demonstrated a therapeutic effect in this model at all doses but with a potentially greater effect of armodafinil of doses of 200 to 300 mg/day. This reviewer generally agrees with the selection of the doses studied in the development program (150 and 250 mg/day). However, it may have been useful to investigate lower doses as well.

Inclusion/Exclusion criteria were satisfactory. Through the use of a variety of criteria, including ESS, CGI-S and sleep latency patients were required to be suffering from a moderate to higher degree of sleepiness. They were similar to the criteria used in the studies with Provigil. Off note, patients who participated in the OSAHS studies were required to suffer residual sleepiness despite the proof that CPAP was effective and patients continued its use on a regular basis (at least 4 hours a night on at least 70% of the nights).

1.17.4 Efficacy Findings

1.17.4.1 OSAHS Studies

Two studies (3025 and 3021) were performed that examined the potential efficacy of armodafinil in OSAHS.

1.17.4.1.1 Demographics and Baseline Characteristics

1.17.4.1.1.1 Protocol 3025

A table describing many of the demographic variables monitored for protocol 3025 is presented in the table below. As would be expected for the general baseline profile for OSAHS were as expected for such a group: i.e. male patients over age 40 years who were over weight (mean BMI 36.5). Demographic features were well matched between placebo and drug treatment groups.

Appears This Way
On Original

Demographic variable Statistic	Armodafinil			p-value
	150 mg (N=129)	Placebo (N=130)	Total (N=259)	
Age, years				
Mean	50.7	50.6	50.7	0.9032 ^a
SD	9.17	8.85	8.99	
Median	52.0	52.0	52.0	
Min, max years	27.0, 69.0 ^b	25.0, 68.0 ^b	25.0, 69.0	
Age group, n (%)				
18-29	1 (<1)	4 (3)	5 (2)	0.5674 ^c
30-40	20 (16)	16 (12)	36 (14)	
41-55	65 (50)	67 (52)	132 (51)	
>55	43 (33)	43 (33)	86 (33)	
Sex, n (%)				
Male	97 (75)	93 (72)	190 (73)	0.5058 ^d
Female	32 (25)	37 (28)	69 (27)	
Race, n (%)				
White	107 (83)	111 (85)	218 (84)	0.4461 ^c
Black	8 (6)	11 (8)	19 (7)	
Asian	2 (2)	1 (<1)	3 (1)	
American Indian or Alaskan Native	0	0	0	
Pacific Islander	0	0	0	
Other	12 (9)	6 (5)	18 (7)	
Missing	0	1 (<1)	1 (<1)	
Race group, n (%)				
White	107 (83)	111 (85)	218 (84)	0.4914 ^d
Nonwhite	22 (17)	18 (14)	40 (15)	
Missing	0	1 (<1)	1 (<1)	
Weight, kg				
Mean	110.7	110.6	110.6	0.9762 ^a
SD	24.04	23.54	23.74	
Median	106.6	107.8	107.6	
Min, max	65.8, 181.4	63.5, 181.9	63.5, 181.9	

Mean ESS scores between placebo and drug groups were similar and indicated a relatively significant degree of sleepiness (16.0 and 15.6, respectively). All patients had a greater than moderately ill rating on the CGI-S with about 40% being markedly ill or worse. CGI-S scores were similar between drug and placebo groups. Mean MWT between experimental groups were also similar (placebo 23.3 min and drug 23.7 min).

Medical histories, based upon organ systems, were similar between placebo and drug treatment groups. The most common organ system reported was that of cardiovascular (60% in drug and 58% in placebo) with hypertension (40% drug and 39 % placebo) constituting the predominant specific disordered. General categories of medications historically used exhibited similar use between both experimental groups with non-steroidal anti-inflammatory agents being most commonly used (45% in drug and 39% in placebo). Antihypertensive drugs were the second most commonly used class (31 % in drug and 38% in placebo).

1.17.4.1.1.2 Protocol 3021

Demographics for the different study groups (safety analysis) are presented in the table below. Demographic features were generally well matched between all the groups. There was a slightly greater number of patients >55 year old and smaller number between 41-40 years old in the placebo group as compared to either drug group. Although not shown mean BMI was well matched between groups (placebo 37.0 kg/m², combined drug group 36.5 kg/m²). As in the 3025 protocol the general population of patients reflected what is known of the demographics of the disease; i.e. the majority of patients were middle age and overweight men.

Demographic variable Statistic	Armodafinil 250 mg/day (N=131)	Armodafinil 150 mg/day (N=131)	Armodafinil combined (N=262)	Placebo (N=130)	Total (N=392)	p-value
Age, years						
Mean	49.1	49.3	49.2	50.1	49.5	0.6551 ^a
SD	8.74	9.17	8.94	9.43	9.10	
Median	50.0	50.0	50.0	52.0	50.0	
Min, max	27.0, 67.0	26.0, 64.0	26.0, 67.0	27.0, 66.0	26.0, 67.0	
Age group (years), n (%)						
18-29	2 (2)	3 (2)	5 (2)	3 (2)	8 (2)	0.6449 ^b
30-40	19 (15)	21 (16)	40 (15)	21 (16)	61 (16)	
41-55	76 (58)	66 (50)	142 (54)	60 (46)	202 (52)	
>55 ^c	34 (26)	41 (31)	75 (29)	46 (35)	121 (31)	
Sex, n (%)						
Male	89 (68)	97 (74)	186 (71)	90 (69)	276 (70)	0.5216 ^d
Female	42 (32)	34 (26)	76 (29)	40 (31)	116 (30)	
Race, n (%)						
White	111 (85)	109 (83)	220 (84)	113 (87)	333 (85)	0.8064 ^e
Black	13 (10)	15 (11)	28 (11)	10 (8)	38 (10)	
Asian	1 (<1)	3 (2)	4 (2)	2 (2)	6 (2)	
American Indian or Alaskan Native	0	0	0	0	0	
Pacific Islander	0	1 (<1)	1 (<1)	0	1 (<1)	
Other	6 (5)	3 (2)	9 (3)	5 (4)	14 (4)	
Race group, n (%)						
White	111 (85)	109 (83)	220 (84)	113 (87)	333 (85)	0.7004 ^f
Nonwhite	20 (15)	22 (17)	42 (16)	17 (13)	59 (15)	

Baseline MWT and CGI-S values were well matched between both treatment groups: e.g. MWT baseline for combined dosage armodafinil and placebo were 22.4 and 23.2 min respectively. CGI-S data are presented in the table below.

Baseline variable Statistic	Armodafinil 250 mg (N=131)	Armodafinil 150 mg (N=131)	Armodafinil combined (N=262)	Placebo (N=130)	Total (N=392)	p-value
CGI-S ratings, n (%) (safety analysis set)						
Normal/not at all ill	0	0	0	0	0	0.6635
Borderline ill	0	0	0	0	0	
Slightly ill	0	0	0	0	0	
Moderately ill	69 (53)	78 (60)	147 (56)	63 (48)	210 (54)	
Markedly ill	42 (32)	34 (26)	76 (29)	43 (33)	119 (30)	
Severely ill	17 (13)	17 (13)	34 (13)	22 (17)	56 (14)	
Among the most extremely ill	3 (2)	2 (2)	5 (2)	2 (2)	7 (2)	

The most common medical history was categorized under HEENT with similar rates of reporting in placebo and drug treated groups (58% in placebo and 59% in drug). The second most, system categorized, disease reported were that of cardiovascular with similar rates observed between groups (57% drug and 59% placebo). Hypertension constituted the largest group in the latter category. Like study 3025 NSAIDs were the most common class of drugs historically used by patients (44% drug 47 % placebo). Antihypertensives were the second most common with 35% on drug and 42% on placebo having used this class of medication. Except for minor differences between groups (as exemplified by the NSAIDS and antihypertensive) drug class use was similar between groups.

1.17.4.1.2 Efficacy Results

1.17.4.1.2.1 Protocol 3025:

1.17.4.1.2.1.1 Primary Endpoints

1.17.4.1.2.1.1.1 MWT

The table below presents the change from baseline the final endpoint evaluation of both experimental groups for the original analysis as well as all post-hoc analyses. Values are in minutes. The table was derived from the FDA statistical review. Evaluations are enumerated according the section on primary endpoints describing **post-hoc analyses**. **As noted in the FDA's** statistical review an alternative ANOVA model was applied because baseline interaction was significant in the primary ANCOVA model. The original NDA database analysis demonstrated a of 3.6 minute prolongation in latency in the armodafinil group as compared to placebo. This was significant at a p value of 0.0003. Non-worst case scenario post-hoc analysis (analyses 1 to 6) demonstrated a similar magnitude of effect (2.9 to 4.3 minute prolongation over placebo) and similarly significant p values (<0.0001 to 0.0037). The worst case scenarios, 7.1 and 7.2,

evaluations demonstrated significance only for 7.2 but with a smaller magnitude than that seen in other post-hoc analyses. Nonetheless, both analyses revealed a trend in the correct direction with prolongations in the armodafinil drug group of 1.6 and 2.2 minutes over placebo. The lack of significance of the 7.2 analysis is not unexpected as the replacement pool for flawed data included data from the evening sessions (1700 and 1900) when drug levels, and therefore effect, may be expected to be lower. The conservative worst case scenario demonstrated no appreciable effect.

Method *	Armodafinil 150 mg N Mean (SD) p-value	Placebo Mean (SD)	treatment difference 95% CI
Analysis 1	116 2.8 (7.67) <0.0001	120 -1.5 (6.39)	2.50, 6.06
Analysis 2	116 1.7 (7.87) 0.0037	120 -1.2 (7.11)	0.93, 4.75
Analysis 3	116 2.4 (7.84) 0.0002	120 -1.3 (7.06)	1.74, 5.56
Analysis 4	116 2.6 (7.61) 0.0002	120 -0.9 (6.81)	1.65, 5.31
Analysis 5	78 2.7 (7.71) 0.0002	73 -1.6 (6.20)	2.08, 6.57
Analysis 6	111 2.6 (8.16) 0.0007	112 -1.0 (7.29)	1.49, 5.51
Analysis 7.1	116 1.7 (8.33) 0.1418	120 0.1 (7.18)	(-0.50, 3.47)
Analysis 7.2	116 2.1 (7.98) 0.0315	120 -0.1 (7.14)	(0.19, 4.00)
Analysis 8.1	116 1.0 (8.34) 0.9285	120 1.0 (7.71)	(-2.15, 1.97)
Analysis 8.2	116 1.0 (8.18) 0.9959	120 1.0 (7.72)	(-2.04, 2.03)
original NDA	116 2.3 (7.80) 0.0003	120 -1.3 (7.08)	1.67, 5.46

Because of the model residuals showed evidence of nonnormality, a Wilcoxon rank-sum test was performed using the original NDA dataset to assess the robustness of the inference based on the ANOVA model fit. The Sponsor noted that inferences from both parametric (ANOVA) and nonparametric (Wilcoxon rank-sum test) tests were consistent in that they both demonstrated significant p-values (<0.05). The Sponsor notes that in accordance with the statistical analysis plan, all efficacy results and the discussion in this report are based on the ANOVA procedures.

Change in CGI-C was examined by using a chi-square test adjusted for country to determine the significance of at least a minimal **improvement in the patient's** condition. Seventy-one percent of patients on drug exhibited at least minimal improvement as compared to 53% of patients on placebo. This was statistically significant at a p value of 0.0069. A breakdown of actual scores for the CGI-C for both experimental groups is presented in the table below. Much of the difference was found in the number of patients in the very much improved group. Fewer percent of patients on armodafinil than placebo exhibited worsening.

CGI-C rating	Number (%) of patients	
	Armodafinil 150 mg (N=116)	Placebo (N=120)
Very much improved	24 (21)	7 (6)
Much improved	30 (26)	26 (22)
Minimally improved	28 (24)	31 (26)
No change	33 (28)	51 (43)
Minimally worse	0	3 (3)
Much worse	1 (<1)	2 (2)
Very much worse	0	0

1.17.4.1.2.1.2 Secondary Endpoints Related to Sleepiness and Fatigue

- MWT (using original database analysis) was assessed at week 4, 8 and 12 as a secondary endpoint averaged for times used in the primary endpoint (0400, 1100, 1300, and 1500). The results for the full analysis set can be found in the table below. As is apparent a statistically significant prolongation in the latency was observed in the drug treated groups as compared to the group receiving placebo at each assessment. There was no obvious time dependency (tolerance) for this effect.

	Week 4		Week 8		Week 12	
	Armodafinil (n=110)	Placebo (n=116)	Armodafinil (n=109)	Placebo (n=108)	Armodafinil (n=108)	Placebo (n=110)
Mean Change from baseline (min)	1.9	-1.1	2.5	-0.3	2.6	-1.6
p value	0.014		0.0039		<0.0001	

- MWT, assessed at weeks 4, 8 and 12, at latter times for the full analysis set (average of three naps at 1500, 1700 and 1900), are presented in the table below. Numerical lengthening of sleep latency was apparent for these latter times but less in magnitude than

the earlier time points. This change in latency, however, was only noted to be significant at week 12.

	Week 4		Week 8		Week 12	
	Armodafinil (n=109)	Placebo (n=113)	Armodafinil (n=109)	Placebo (n=108)	Armodafinil (n=107)	Placebo (n=107)
Mean Change from baseline (min)	1.6	-0.1	2.2	0.6	1.8	-0.5
P value	0.0717		0.1004		0.0435	

- Number (and percent) of patients with at least a minimal improvement in the CGI-C at weeks 4 8 and 12 for the full analysis set are presented in the table below. There was a significant improvement in symptoms based upon this metric at every time point, although, numerically, the effect may have been minimally more pronounced at week 4. The magnitude of this difference was not large enough to suggest tolerance over this period of time. Similar to the primary endpoint for CGI-C, this effect was primarily expressed as increased number of patients who were reported to **have a rating of “much improved” and “very much improved” (data not shown).**

Time point, variable	Number (%) of patients	
	Armodafinil 150 mg (N=116)	Placebo (N=120)
	Week 4, n (%)	112 (100)
At least minimal improvement ^a	86 (77)	54 (47)
p-value	<0.0001	
Week 8, n (%)	111 (100)	110 (100)
At least minimal improvement ^a	82 (74)	57 (52)
p-value	0.0008	
Week 12, n (%)	110 (100)	110 (100)
At least minimal improvement ^a	76 (69)	60 (55)
p-value	0.0282	

- The baseline, endpoint difference mean ESS scores for the full analysis set for placebo and drug treated groups is presented in the table below for endpoint and each week analysis. Both group placebo and drug groups exhibited a reduction in sleepiness. The reduction was greater for the armodafinil group. This effect was statistically significant at every time point and sufficient to take the patient from a high degree of moderate sleepiness to a low degree of moderate sleepiness or mild sleepiness. Thus, there were no obvious signs of drug tolerance over the period of time studied.

	Armodafinil (n=116)	Placebo (n=120)	p-value
--	---------------------	-----------------	---------

Baseline	15.6	16.0	
Mean Change from baseline at 4 Weeks	-4.6	-3.0	0.0069
Mean Change from baseline at 8 weeks	-5.3	-3.0	0.0002
Mean Change from baseline at 12 weeks	-5.2	-2.9	0.0004
Mean Change from baseline at endpoint	-5.3	-3.0	0.0001

- Mean baseline values and changes in the mean BFI from baseline to the endpoint for the full analysis set for placebo and combined armodafinil dose groups are presented in the table below. Note there was a significant reduction in the BFI in the drug as compared to the placebo group, presumably indicating a greater reduction in fatigue. As this metric has not been studied in sleep disorders it is difficult to determine the meaning of such a change. It is, however, in the appropriate direction. Moreover, examination of weeks 2, 8 and 12 endpoints revealed similar differences. These, however, were only significantly different for weeks 4 and 12.

	Armodafinil (n=115)	Placebo (n=119)
Mean BFI	4.7	4.9
Endpoint change from baseline	-1.3	-0.4
P value	0.0184	

- The table below presents information from daytime diaries. Data were not rigorously statistically analyzed but all changes observed were in the direction of a therapeutic effect. Thus, there was a greater decrease in the number of unintentional episodes of daily sleep and naps in the armodafinil then the placebo group. Endpoint mistakes, accidents and near misses were less in the drug treated group. Although baseline caffeine intake was somewhat higher in the placebo group there was little difference in caffeine intake following both treatments.

		Armodafinil	Placebo
Unintentional Sleep episodes	Mean episodes at baseline	1.0	1.2
	Mean % change from baseline at endpoint	-54.7%	-37.1%
Number of daily naps	Mean number of naps at baseline	0.5	0.5
	Mean % change from baseline at endpoint	-35.7%	-17.0%
Percent of patients reporting mistakes Accidents and near misses			
	At endpoint	78%	83%

Clinical Review
Norman Hershkowitz, MD, PhD
21,875 (000)
Nuvigil (armodafinil)

Caffeine beverage use	Mean number at baseline	2.1	2.8
	Mean number change	-0.2	-0.2

1.17.4.1.2.1.3 Secondary Endpoints Related to Cognition (CDR)

1.17.4.1.2.1.3.1 Key Secondary Endpoint- Power of Attention

A key secondary endpoint was to determine the effect of armodafinil on attention as measured by the change from baseline in the mean power of attention from the CDR system (mean of 4 evaluations 0930 to 1530). Analysis was performed with the use of an ANCOVA with treatment and country as factors and baseline as a covariate. If the final model for the primary efficacy analysis used an ANOVA instead of an ANCOVA, then the key secondary analysis used an ANOVA. No separate treatment by covariate interaction or treatment by country interaction was tested for the key secondary analysis. Note the final test for significance in this CDR measure and all subsequent ones for this study was an ANOVA. Results of this analysis at the final endpoint can be found in the table below (values are in msec). There were very minor increases in reaction time following placebo and drug. This effect was greatest for drug, but to a very small degree (about 5 msec). These differences were not found to be statistically significant (ANOVA).

Appears This Way
On Original

Time point^a Statistic	Armodafinil 150 mg (N=116)	Placebo (N=120)
Baseline		
n	116	120
Mean	1251.9	1308.9
SD	140.22	190.02
Median	1230.5	1274.0
Min, max	1011.7, 1786.9	977.8, 2304.0
Endpoint		
n	116	120
Mean	1300.4	1352.5
SD	175.98	215.65
Median	1264.1	1327.3
Min, max	1008.9, 2078.7	994.3, 2924.1
Change from baseline to endpoint		
n	116	120
Mean	48.6	43.6
SD	87.21	208.81
Median	42.5	38.0
Min, max	-136.1, 317.4	-912.3, 1673.6
p-value	0.8181	

1.17.4.1.2.1.3.2 Other Cognitive CDR Testing

- The results of the continuity of attention for the endpoint evaluation (average of 4 tests at times 0930, 1130, 1330, and 1530) for the full analysis set at the final endpoint are presented in the table below. No significant difference can be appreciated between drug and placebo groups. There were no meaningful or statistically significant difference between both groups (ANOVA).

Appears This Way
 On Original

Time point^a Statistic	Armodafinil 150 mg (N=116)	Placebo (N=120)
Baseline		
n	116	120
Mean	91.1	90.2
SD	2.70	4.65
Median	91.6	91.3
Min, max	79.3, 94.5	58.8, 94.8
Endpoint		
n	116	120
Mean	91.1	89.8
SD	3.44	4.18
Median	91.8	91.0
Min, max	74.2, 95.0	70.7, 94.5
Change from baseline to endpoint		
n	116	120
Mean	0.0	-0.3
SD	2.37	4.14
Median	0.0	-0.2
Min, max	-10.3, 6.8	-18.5, 25.3
p-value	0.4477	

- The quality of episodic secondary memory from the CDR (average of 4 tests at 0930, 1130, 1330 and 1530) for full analysis set at the final endpoint evaluation is presented in the table below. Armodafinil produced a statistically significant improvement in this score. The Sponsor also examined these indices at latter times (1530, 1730 and 1930). Although there was a difference in favor of armodafinil at the endpoint measure it was not statistically significant (ANOVA).

Appears This Way
 On Original

Clinical Review
 Norman Hershkowitz, MD, PhD
 21,875 (000)
 Nuvigil (armodafinil)

Time point^a Statistic	Armodafinil 150 mg (N=116)	Placebo (N=120)
Baseline		
n	116	120
Mean	172.4	161.9
SD	40.30	44.10
Median	172.1	162.3
Min, max	65.8, 287.9	58.8, 281.7
Endpoint		
n	116	120
Mean	180.1	154.9
SD	42.72	56.46
Median	177.1	160.8
Min, max	49.6, 267.5	-134.2, 265.0
Change from baseline to endpoint		
n	116	120
Mean	7.6	-7.0
SD	30.66	52.58
Median	6.0	-5.6
Min, max	-83.7, 75.8	-415.8, 91.2
p-value	0.0102	

Appears This Way
 On Original

- The Speed of memory (average of 4 tests at 0930, 1130, 1330, and 1530) from the full analysis set at the final endpoint is presented in the table below. No difference between placebo and armodafinil treated groups were appreciated (ANOVA).

Time point ^a Statistic	Armodafinil 150 mg (N=116)	Placebo (N=120)
Baseline		
n	116	120
Mean	2818.8	2887.4
SD	490.73	616.78
Median	2821.4	2768.2
Min, max	1952.2, 4389.4	1904.6, 5734.4
Endpoint		
n	116	120
Mean	2618.5	2685.0
SD	452.50	511.77
Median	2579.2	2593.3
Min, max	1800.3, 3967.2	1811.1, 4926.6
Change from baseline to endpoint		
n	116	120
Mean	-200.3	-202.4
SD	278.34	345.29
Median	-209.8	-153.6
Min, max	-989.9, 850.1	-1529.8, 648.0
p-value	0.8686	

1.17.4.1.2.2 Protocol 3021:

1.17.4.1.2.2.1 Primary Endpoints

1.17.4.1.2.2.1.1 MWT

The table below presents the change from baseline to the final endpoint evaluation of all experimental groups for the original analysis as well as the post-hoc analyses. The table was derived from the FDA statistical review. **Values are in minutes. As noted in the FDA's statistical review, an alternative ANOVA model was applied because baseline interaction was**

significant in the primary ANCOVA model. The original NDA database analysis demonstrated a of 3.6 minute prolongation in latency in the armodafinil combined dose group as compared to placebo. This was significant at a p value of <0.0001. Each individual dose also exhibited similar magnitudes of prolongation and highly significant p values. Based upon planned sequential (high dose to low) this analysis would suggest a therapeutic effect for both doses. Non-worst case scenario post-hoc analysis (analyses 1 to 6) demonstrated a similar magnitude of effect (2.9 to 3.8 minute prolongation over placebo in the combined dose group) and similarly significant p values (<0.0001 to 0.0037). All individual dose evaluations for analyses 1-6 were significant as well. Like the other OSAHS study, the worst case scenarios, 7.1 and 7.2, evaluations demonstrated statistically significance prolongation only for 7.2 but of smaller magnitudes then other post-hoc analyses. All individual dose analyses for the 7.1 were significant. Both worst case scenario analyses (7.1 and 7.2), nonetheless, trended in the correct direction: e.g. combined armodafinil drug group exhibited latency prolongations of 1.7 to 2.2 minutes over placebo. The lack of significance of the 7.2 analysis is not unexpected as the replacement pool for flawed data included data from the evening sessions (1700 and 1900) when drug levels, and therefore effect, may be expected to be lower. The conservative worst case scenario (8.1 and 8.2) did not show a statistically significant effect but trended in the correct direction with the combined dose prolongations over placebo of 1.3 to 1.5 minutes. There was a trend for greater prolongation with the higher armodafinil dose.

Method *	Armodafinil 250 mg	Armodafinil 150 mg	Armodafinil Combined	Placebo	treatment difference 95% CI
	N Mean (SD) p-value	N Mean (SD) p- value	N Mean (SD) p-value	N Mean (SD)	
Analysis 1	121 2.3 (8.02) 0.0005	120 2.1 (6.70) 0.0007	241 2.2 (7.38) <0.0001	124 -1.2 (8.19)	1.75, 5.08
Analysis 2	121 1.8 (7.81) 0.0015	120 1.5 (6.49) 0.0050	241 1.6 (7.17) 0.0006	124 -1.3 (8.50)	1.28, 4.61
Analysis 3	121 2.5 (7.97) 0.0001	120 1.8 (6.90) 0.0014	241 2.1 (7.45) <0.0001	124 -1.3 (8.12)	1.82, 5.16
Analysis 4	121 2.4 (7.84) 0.0001	120 2.0 (6.39) 0.0006	241 2.2 (7.14) <0.0001	124 -1.4 (8.36)	1.94, 5.33
Analysis 5	84 2.6 (8.00) 0.0011	90 1.8 (6.05) 0.0088	164 2.2 (7.11) 0.0006	92 -1.2 (8.25)	1.47, 5.34
Analysis 6	118 2.6 (8.19) <0.0001	120 1.8 (7.41) 0.0013	238 2.2 (7.80) <0.0001	123 -1.6 (8.59)	2.01, 5.55
Analysis 7.1	121 1.5 (8.35) 0.0613	120 1.0 (7.25) 0.1636	241 1.2 (7.81) 0.0590	124 -0.5 (8.44)	(-0.06, 3.43)
Analysis 7.2	121 1.6 (8.35) 0.0258	120 1.3 (7.35) 0.0479	241 1.5 (7.85) 0.0150	124 -0.7 (8.29)	(0.43, 3.90)
Analysis 8.1	121 1.1 (8.58) 0.1423	120 0.8 (7.33) 0.2551	241 0.9 (7.97) 0.1317	124 -0.4 (8.42)	(-0.41, 3.12)
Analysis 8.2	121 1.1 (8.58) 0.1178	120 0.8 (7.37) 0.2174	241 1.0 (7.99) 0.1056	124 -0.5 (8.38)	(-0.31, 3.23)
Original NDA	121 2.3 (8.07) 0.0001	120 1.7 (6.49) 0.0006	241 1.9 (7.32) <0.0001	124 -1.7 (8.59)	1.93, 5.31

Appears This Way
 On Original

Because of the model residuals showed evidence of nonnormality, a Wilcoxon rank-sum test was performed using the original NDA dataset to assess the robustness of the inference based on the ANOVA model fit. The Sponsor noted that inferences from both parametric (ANOVA) and nonparametric (Wilcoxon rank-sum test) tests were consistent in that they both demonstrated significant p-values (<0.05). The Sponsor notes that in accordance with the statistical analysis plan, all subsequent efficacy analysis (i.e. secondary endpoints) and the discussion is based on the ANOVA procedures.

1.17.4.1.2.2.1.2 CGI-C

The primary endpoint, comparing patients with at least a minimal improvement revealed efficacy. Thus 74% and 71% of patients in high and low dose drug groups, respectively, experienced at least minimal improvement as compared to 37% of patients on placebo. This was statistically significant (p=0.0001, Cochran-Mantel-Haenzel chi square) for all comparisons including combined and individual doses sequentially examined. The differences between doses were probably clinically insignificant.

The table below presents changes in the CGI-C for the various experimental groups. The data suggests efficacy by the fact that many more patients on drug than on placebo showed very much improved or much improved. Fewer percent of patients on armodafinil than placebo exhibited worsening.

CGI-C rating	Number (%) of patients			
	Armodafinil 250 mg/day (N=121)	Armodafinil 150 mg/day (N=120)	Armodafinil combined (N=241)	Placebo (N=124)
Very much improved	28 (23)	30 (25)	58 (24)	10 (8)
Much improved	36 (30)	40 (33)	76 (32)	18 (15)
Minimally improved	25 (21)	15 (13)	40 (17)	18 (15)
No change	28 (23)	34 (28)	62 (26)	71 (57)
Minimally worse	3 (2)	1 (<1)	4 (2)	6 (5)
Much worse	1 (<1)	0	1 (<1)	1 (<1)
Very much worse	0	0	0	0

1.17.4.1.2.2.2 Secondary Endpoints Related to Sleepiness and Fatigue

- MWT (using original NDA datasets) was assessed at week 4, 8 and 12 as a secondary endpoint averaged for times used in the primary endpoint (0400, 1100, 1300, and 1500). The results for the full analysis set can be found in the table below. As is apparent a statistically significant prolongation (ANOVA) in the latency was observed in the drug

treated groups as compared to the group receiving placebo. High and low dose analysis was also observed to be statistically significantly different from placebo at each time point. There was no obvious time dependency or tolerance of this effect.

	Week 4		Week 8		Week 12	
	Armodafinil (combined groups) (n=241)	Placebo (n=124)	Armodafinil (combined groups) (n=216)	Placebo (n=116)	Armodafinil (combined groups) (n=210)	Placebo (n=118)
Mean Change from baseline (min)	2.3	0.0	1.7	-0.1	1.8	-1.3
P value	0.0047		0.0344		0.0004	

- MWT, accessed at weeks 4, 8 and 12, at latter times for the full analysis set (average of three naps at 1500, 1700 and 1900), are presented in the table below. Numerical lengthening of sleep latency was suggested by mean changes for these latter times but of less in magnitude than the earlier time points. This change in latency, however, was only noted to be significant at week 4.

	Week 4		Week 8		Week 12	
	Armodafinil (combined groups) (n=239)	Placebo (n=124)	Armodafinil (combined groups) (n=213)	Placebo (n=116)	Armodafinil (combined groups) (n=209)	Placebo (n=117)
Mean Change from baseline (min)	-0.3	1.5	0.7	0.3	0.8	-0.1
P value	0.015		0.6234		0.3166	

- Number (and percent) of patients with at least a minimal improvement in the CGI-C at weeks 4, 8 and 12 for the full analysis set are presented in the table below. There was a statistically significant improvement in symptoms (Cochran-Mantel-Haenzel chi square) for combined doses and all doses for all times. There is no evidence of drug tolerance as measured by this endpoint. Except for a sequential analysis plan (first combined and then individual doses) the analysis is not corrected for multiple comparisons. The magnitude of effect was not substantially different at the different time points.

Appears This Way
 On Original

CGI-C rating	Number (%) of patients			
	Armodafinil 250 mg/day (N=121)	Armodafinil 150 mg/day (N=120)	Armodafinil combined (N=241)	Placebo (N=124)
Week 4, n (%)	115 (100)	116 (100)	231 (100)	121 (100)
At least minimal improvement ^a	87 (76)	79 (68)	166 (72)	47 (39)
p-value	<0.0001	<0.0001	<0.0001	—
Week 8, n (%)	110 (100)	114 (100)	224 (100)	116 (100)
At least minimal improvement ^a	85 (77)	87 (76)	172 (77)	47 (41)
p-value	<0.0001	<0.0001	<0.0001	—
Week 12, n (%)	108 (100)	113 (100)	221 (100)	120 (100)
At least minimal improvement ^a	83 (77)	82 (73)	165 (75)	46 (38)
p-value	<0.0001	<0.0001	<0.0001	—

- The mean ESS baseline score and mean and median differences from the baseline at different time points as well as the endpoint measure for the full analysis set for placebo and combined drug dose groups is presented in the table below. Both placebo group and drug groups exhibited a reduction in sleepiness. The reduction was greater for the armodafinil group. This effect was statistically significant at every time point (not corrected for multiple comparisons). Thus, there was no evidence of tolerance of effect as measured by this endpoint. Individual dose groups (data not shown) exhibited similar reduction in sleepiness as measured by the **ESS and each dose group's reduction in sleepiness** was statistically significant with the highest p value of 0.0007 (not corrected for multiple comparisons).

		Combined Armodafinil (n=239)	Placebo (n=123)	p-value
	Baseline	15.3	15.9	
Week 4	Mean Change from baseline	-5.2	-2.7	<0.0001
	Median Change from baseline	-5.0	-2.0	
Week 8	Mean Change from baseline	-5.7	-3.2	<0.0001
	Median Change from baseline	-5.0	-2.0	
Week 12	Mean Change from baseline	-5.7	-3.3	<0.0001

	Median Change from baseline	-5.0	-3.0	
Endpoint	Mean Change from baseline	-5.5	-3.3	<0.0001
	Median Change from baseline	-5.0	-3.0	

- Mean baseline values and changes in the mean BFI from baseline to the endpoint for the full analysis set for placebo and combined armodafinil dose groups are presented in the table below. Note there was a significant reduction in the BFI score, presumably indicating a reduction in fatigue. As noted above, the meaning of this is not clear as this metric has not been validated for this population of patients. Each individual dose was also noted to have a statistically significant reduction (data not shown). Moreover, examination of weeks 2, 8 and 12 endpoints revealed a similar statistically significant difference between placebo and the combined dose groups.

	Combined Armodafinil (n=115)	Placebo (n=119)
Mean BFI	5.2	4.6
Endpoint change from baseline	-1.2	-0.4
P value	0.0059	

- The table below presents information from daytime diaries. There was a greater decrease in the number of unintentional sleep episodes and daily naps in the armodafinil then the placebo group. Accidents and mistakes were similar across groups. Baseline caffeine use was similar in both groups with little difference between groups following treatment.

		Combined Armodafinil	Placebo
Unintentional Sleep episodes	Mean episodes at baseline	1.0	1.0
	Mean % change from baseline at endpoint	-55.5%	-18.9%
Number of daily naps	Mean number of naps at baseline	0.4	0.4
	Mean % change from baseline at endpoint	-34.4%	-3.5%
Percent of patients reporting mistakes Accidents and near misses	At endpoint	77%	77%
Caffeine beverage use	Mean number at baseline	2.4	2.2
	Mean number change	-0.3	-0.1

1.17.4.1.2.2.3 Secondary Endpoints Related to Cognition (CDR)

1.17.4.1.2.2.3.1 *Key Secondary Endpoint—Quality of Episodic Memory*

The quality of episodic secondary memory from the CDR (average of 4 tests at 0930, 1130, 1330 and 1530) for full analysis set at the final endpoint evaluation was not found to be statistically significant (ANOVA). The table below presents this data.

(Full Analysis Set)

Time point^a Statistic	Armodafinil 250 mg/day (N=121)	Armodafinil 150 mg/day (N=120)	Armodafinil combined (N=241)	Placebo (N=124)
Baseline				
n	121	119	240	123
Mean	171.9	169.4	170.6	174.9
SD	39.25	46.05	42.69	40.22
Median	169.6	173.8	171.2	171.3
Min, max	99.2, 270.0	27.1, 271.7	27.1, 271.7	80.0, 342.9
Endpoint				
n	121	119	240	124
Mean	182.8	181.3	182.1	180.5
SD	45.70	50.57	48.08	53.19
Median	182.1	185.8	184.0	182.7
Min, max	60.8, 286.7	-29.6, 334.2	-29.6, 334.2	-52.5, 350.8
Change from baseline to endpoint				
n	121	119	240	123
Mean	10.9	11.9	11.4	5.4
SD	30.97	33.57	32.22	38.36
Median	8.7	13.8	11.2	2.9
Min, max	-72.1, 133.3	-92.9, 102.1	-92.9, 133.3	-242.5, 86.7
p-value	—	—	0.1147	—

1.17.4.1.2.2.3.2 *Other Cognitive CDR Testing*

- The results of the continuity of attention for the endpoint evaluation (average of 4 tests at times 0930, 1130, 1330, and 1530) for the full analysis was not observed to be statistically significant between the various groups (ANOVA) at the final endpoint.
- The Speed of memory (average of 4 tests at 0930, 1130, 1330, and 1530) from the full analysis set at the final endpoint was not found to be statistically significant between the experimental groups (ANOVA).
- Data for the mean power of attention at the final endpoint from the CDR system (mean of 4 evaluations 0930 to 1530) are presented in the table below. Analysis is similar to that described above and are presented in the table below with values in terms of msec. There were minor increases in reaction times for both drug and placebo groups. This increase was slightly greater in the placebo group. This difference was not found to be statistically significant when comparing the combined drug group to placebo (ANOVA).

Time point* Statistic	Armodafinil 250 mg/day (N=121)	Armodafinil 150 mg/day (N=120)	Armodafinil combined (N=241)	Placebo (N=124)
Baseline				
n	121	119	240	123
Mean	1266.8	1250.1	1258.5	1256.8
SD	195.17	166.93	181.53	149.50
Median	1225.8	1231.0	1228.4	1252.6
Min, max	1039.6, 2443.3	976.4, 2132.8	976.4, 2443.3	1009.0, 1752.0
Endpoint				
n	121	119	240	124
Mean	1299.0	1287.6	1293.4	1306.4
SD	257.72	180.24	222.31	171.02
Median	1252.2	1261.1	1253.3	1305.9
Min, max	1047.3, 3548.6	1011.2, 2041.7	1011.2, 3548.6	1010.4, 1770.7
Change from baseline to endpoint				
n	121	119	240	123
Mean	32.3	37.5	34.9	48.0
SD	210.47	96.75	163.92	99.94
Median	19.0	29.2	25.0	43.9
Min, max	-1164.9, 1702.1	-388.6, 381.5	-1164.9, 1702.1	-301.0, 351.9
p-value	—	—	0.4170	—

Appears This Way
 On Original

1.17.4.2 Narcolepsy Study (3020)

1.17.4.2.1 Demographics and Baseline Characteristics

The table below presents demographic information on the studied patient population for this study. There were small disparities in the age between drug and placebo. The numerical age differences were, however, not of great magnitude. A slightly disproportionate number of males and patients of the “white” race were observed in the placebo group. It is impossible to state with certainty how these may influence data, but this reviewer feels that the effect is likely minimal.

Demographic variable Statistic	Armodafinil 250 mg (N=67)	Armodafinil 150 mg (N=64)	Armodafinil combined (N=131)	Placebo (N=63)	Total (N=194)	p-value
Age, years						
n	67	64	131	63	194	
Mean	35.0	40.4	37.7	39.2	38.1	0.0355 ^a
SD	12.52	12.52	12.76	11.98	12.50	
Median	32.0	43.0	36.0	36.0	36.0	
Min, max	18.0, 67.0	18.0, 65.0	18.0, 67.0	20.0, 63.0	18.0, 67.0	
Age group (years), n (%)						
18 - 29	30 (45)	18 (28)	48 (37)	15 (24)	63 (32)	0.0226 ^b
30 - 40	20 (30)	12 (19)	32 (24)	19 (30)	51 (26)	
41 - 55	11 (16)	27 (42)	38 (29)	22 (35)	60 (31)	
>55 ^c	6 (9)	7 (11)	13 (10)	7 (11)	20 (10)	
Sex, n (%)						
Male	25 (37)	28 (44)	53 (40)	32 (51)	85 (44)	0.3017 ^b
Female	42 (63)	36 (56)	78 (60)	31 (49)	109 (56)	
Race, n (%)						
White	48 (72)	44 (69)	92 (70)	49 (78)	141 (73)	0.7737 ^d
Black	9 (13)	13 (20)	22 (17)	10 (16)	32 (16)	
Asian	1 (1)	1 (2)	2 (2)	0	2 (1)	
American Indian or Alaskan Native	2 (3)	1 (2)	3 (2)	1 (2)	4 (2)	
Pacific Islander	1 (1)	1 (2)	2 (2)	0	2 (1)	
Other	3 (4)	2 (3)	5 (4)	0	5 (3)	
Missing	3 (4)	2 (3)	5 (4)	3 (5)	8 (4)	
BMI, kg/m²						
n	67	64	131	62	193	
Mean	28.3	29.6	29.0	28.3	28.7	0.4120 ^a
SD	6.91	6.71	6.82	5.25	6.35	
Median	28.0	28.9	28.1	28.2	28.2	
Min, max	15.4, 52.7	18.8, 59.2	15.4, 59.2	18.6, 40.6	15.4, 59.2	

SOURCE: Summary 15.2; Listing 5.

^a The p-value for the overall treatment comparison is from an analysis of variance (ANOVA) with treatment group and center as a factor.

^b The p-value for the overall treatment comparison is from a Pearson's chi square test.

^c Patient 3183002 in the armodafinil 250-mg/day treatment group was older than 65 years, but was permitted to enroll in the study.

^d The p-value for the overall treatment comparison is from a Fisher's exact test.

BMI=body mass index; min=minimum; max=maximum.

NOTE: Other=Hispanic (2); Filipino (1); Eurasian (1); Caucasian and Asian (1).

Baseline features sleepiness values for CGI-S and MSLT are presented in the table below. There are no significant differences between placebo and all drug groups. MWT baseline value,

however, was slightly lower in the combined armodafinil group than in placebo. Thus, baseline values of 12.5 minutes and 10.8 minutes were observed in the placebo and combined drug dosage group, respectively. The low armodafinil group was quite similar to the placebo: e.g. low and high dose groups were 12.1 and 9.5 minutes. This can allow for some internal control for the baseline of the test results.

Baseline variable Statistic	Armodafinil 250 mg (N=67)	Armodafinil 150 mg (N=64)	Armodafinil combined (N=131)	Placebo (N=63)	Total (N=194)	p-value
CGI-S ratings, n (%) (safety analysis set)						
Normal-not at all ill	0	0	0	0	0	0.7672*
Borderline ill	0	0	0	0	0	
Slightly ill	0	0	0	0	0	
Moderately ill	25 (37)	19 (30)	44 (34)	18 (29)	62 (32)	
Markedly ill	29 (43)	32 (50)	61 (47)	34 (54)	95 (49)	
Severely ill	12 (18)	11 (17)	23 (18)	11 (17)	34 (18)	
Among the most extremely ill	1 (1)	2 (3)	3 (2)	0	3 (2)	
MSLT sleep latency, minutes (full analysis set)						
n	60	58	118	58	—	—
Mean	2.6	2.5	2.5	2.6	—	—
SD	1.61	1.72	1.66	1.55	—	—
Median	2.4	2.4	2.4	2.3	—	—
Min, max	0.1, 6.1	0.0, 6.6	0.0, 6.6	0.3, 6.0	—	—

SOURCE: Summary 15.3 and section 16.1.9; Listing 10 and Listing 11.

* The p-value for the overall treatment comparison is from a Fisher's exact test.

CGI-S=Clinical Global Impression of Severity; MSLT=Multiple Sleep Latency Test; min=minimum; max=maximum; SD=standard deviation.

Thus, although some indices appeared to differ others were similar between treatment groups. This reviewer does not believe that these differences will markedly affect the final results of the study.

1.17.4.2.2 Efficacy Results

1.17.4.2.2.1 Primary Endpoints

1.17.4.2.2.1.1 MWT

The table below presents the change from baseline to final endpoint evaluation of all experimental groups for the original analysis as well as all post-hoc analyses. The table was derived from the FDA statistical review. Evaluations are enumerated according to the section on primary endpoints describing post-hoc analyses (see above). The change from baseline in sleep latency from the MWT was tested using ANCOVA model with treatment and country as factors, and the corresponding baseline value as a covariate. The original NDA database analysis demonstrated a of 3.8-minute prolongation in latency in the armodafinil combined dose group as compared to placebo. This was significant at a p value of 0.0024. Each individual dose also exhibited similar magnitudes of prolongation and highly significant p values. Based upon

planned sequential (high dose to low) this analysis would suggest a therapeutic effect for both doses. Non-worst case scenario post-hoc analysis (analyses 1 to 6) demonstrated a similar magnitude of effect (4.0 to 4.5 minute prolongation over placebo in the combined dose group) with similarly significant p values (<0.0001 to 0.0072). All individual dose evaluations for analyses 1-6 were significant as well. There was a trend for greater sleep latency prolongation in the higher doses. The worst case scenario (7.1 and 7.2) evaluations demonstrated significance for only the combined and high dose groups in the in 7.2 analyses. As noted above the 7.1 analysis may be considered to bias the data against revealing an effect. Nonetheless, both the 7.1 and 7.2 of combined doses and dose groups trended in the direction toward prolongation. The conservative worst case analysis did not demonstrate significance effect.

Method *	Armodafinil 250 mg N Mean (SD) p-value	Armodafinil 150 mg N Mean (SD) p-value	Armodafinil Combined N Mean (SD) p-value	Placebo N Mean (SD)	treatment difference 95% CI
Analysis 1	60 2.3 (5.84) 0.0002	58 1.3 (5.11) 0.0005	118 1.8 (5.50) <0.0001	58 -2.3 (5.68)	1.84, 5.04
Analysis 2	60 2.0 (6.08) 0.0007	58 1.8 (5.30) 0.0003	118 1.9 (5.69) <0.0001	58 -2.3 (5.85)	1.88, 5.36
Analysis 3	60 2.7 (5.56) 0.0001	58 0.9 (5.32) 0.0017	118 1.8 (5.49) <0.0001	58 -2.2 (5.81)	1.77, 5.04
Analysis 4	60 2.3 (5.72) 0.0001	58 0.8 (4.92) 0.0010	118 1.6 (5.37) <0.0001	58 -2.4 (5.62)	1.75, 4.87
Analysis 5	40 2.6 (6.11) 0.0085	26 1.4 (4.64) 0.0292	66 2.1 (5.57) 0.0072	25 -2.5 (6.76)	1.13, 6.53
Analysis 6	58 2.8 (5.91) 0.0003	54 0.7 (5.12) 0.0050	112 1.8 (5.61) 0.0002	55 -2.2 (5.95)	1.60, 5.06
Analysis 7.1	60 1.9 (5.94) 0.0279	58 -0.7 (5.56) 0.4064	118 0.6 (5.88) 0.0795	58 -1.2 (6.30)	-0.18, 3.24
Analysis 7.2	60 2.1 (5.93) 0.0047	58 -0.3 (5.16) 0.1206	118 0.9 (5.67) 0.0114	58 -1.6 (6.18)	(0.49, 3.86)
Analysis 8.1	60 1.1 (6.52) 0.6716	58 -1.2 (5.56) 0.6176	118 0.0 (6.15) 0.9633	58 0.1 (6.91)	(-1.79, 1.73)
Analysis 8.2	60 1.2 (6.50) 0.5739	58 -1.0 (5.45) 0.7473	118 0.1 (6.08) 0.8911	58 -0.0 (6.89)	(-1.62, 1.88)
original NDA	60 2.6 (6.24) 0.0099	58 1.3 (6.31) 0.0068	118 1.9 (6.28) 0.0024	58 -1.9 (6.87)	1.02, 4.61

Because of the model residuals showed evidence of nonnormality, a Wilcoxon rank-sum test was performed using the original NDA dataset to assess the robustness of the inference based on the ANOVA model fit. The Sponsor noted that inferences from both parametric (ANOVA) and nonparametric (Wilcoxon rank-sum test) tests were consistent in that they both demonstrated significant p-values (<0.05). The Sponsor notes that in accordance with the statistical analysis plan, all efficacy results and the discussion in this report are based on the ANOVA procedures.

1.17.4.2.2.1.2 CGI-C

The primary endpoint, comparing patients with at least a minimal improvement revealed efficacy for combined dose as well as all dose groups ($p=0.0001$ for all comparison-combined and sequential comparison (Cochran-Mantel-Haenzel chi square). This data is presented in the table below. Although, the higher dose group exhibited a greater percent of patients achieving the criteria, it is unknown if this represents a clinically significant greater difference as the study was not designed to examine this issue statistically.

CGI-C rating	Number (%) of patients			
	Armodafinil 250 mg/day (N=60)	Armodafinil 150 mg/day (N=58)	Armodafinil combined (N=118)	Placebo (N=58)
At least minimal improvement	44 (73)	40 (69)	84 (71)	19 (33)
p-value	<0.0001	<0.0001	<0.0001	—

A complete accounting for final CGI-C scores is presented in the table below. From this data it is apparent that much of the difference in affect between placebo and armodafinil originates from patients who exhibited at least a “much improved” or “very much improved score.” A lower percent of patients on armodafinil then placebo worsened.

CGI-C rating	Number (%) of patients			
	Armodafinil 250 mg/day (N=60)	Armodafinil 150 mg/day (N=58)	Armodafinil combined (N=118)	Placebo (N=58)
Very much improved	11 (18)	9 (16)	20 (17)	2 (3)
Much improved	21 (35)	19 (33)	40 (34)	7 (12)
Minimally improved	12 (20)	12 (21)	24 (20)	10 (17)
No change	14 (23)	13 (22)	27 (23)	29 (50)
Minimally worse	1 (2)	3 (5)	4 (3)	4 (7)
Much worse	1 (2)	2 (3)	3 (3)	3 (5)
Very much worse	0	0	0	3 (5)

1.17.4.2.2.2 Secondary Endpoints Related to Sleepiness and Fatigue

- MWT (using the original dataset) were evaluated for weeks 4, 8 and 12. Combined armodafinil doses and placebo values are presented in the table below. All times points were statistically different from placebo (not corrected for multiple comparisons) and there was no obvious numerical evidence for tolerance to this effect over the 12 week period. With the exception of the 250 mg/day dose at 8 and 12 weeks all individual dose differences were statistically significant.

	Week 4		Week 8		Week 12	
	Armodafinil (combined groups) (n=118)	Placebo (n=58)	Armodafinil (combined groups) (n=110)	Placebo (n=55)	Armodafinil (combined groups) (n=88)	Placebo (n=43)
Mean Change from baseline (min)	2.1	-1.1	1.7	-1.3	1.8	-1.7
p value	0.0054		0.0481		0.0264	

- CGI-C endpoint for the different test periods are presented in the table below. All values (each doses and combined dose) where statistically significantly improved (Cochran-Mantel-Haenzel chi square, not corrected for multiplicity) over placebo. There was no obvious time effect (i.e. tolerance) observed.

CGI-C rating	Number (%) of patients			
	Armodafinil 250 mg/day (N=60)	Armodafinil 150 mg/day (N=58)	Armodafinil combined (N=118)	Placebo (N=58)
Week 4, n (%)	57 (100)	57 (100)	114 (100)	56 (100)
At least minimal improvement ^a	40 (70)	43 (75)	83 (73)	22 (39)
p-value	0.0003	<0.0001	<0.0001	—
Week 8, n (%)	58 (100)	55 (100)	113 (100)	53 (100)
At least minimal improvement ^a	41 (71)	37 (67)	78 (69)	20 (38)
p-value	0.0006	0.0024	0.0001	—
Week 12, n (%)	46 (100)	43 (100)	89 (100)	43 (100)
At least minimal improvement ^a	34 (74)	30 (70)	64 (72)	13 (30)
p-value	<0.0001	0.0003	<0.0001	—

- Data for the combined dosage ESS at various times throughout the study and at the final endpoint are presented in the table below. ESS score indicates statistically significant decreased (ANCOVA, not corrected for multiple comparisons) sleepiness with armodafinil at all time points and at the final endpoint evaluation. There does not appear to be any obvious tolerance to the drugs effect of the time period studied. Examination of individual dosages reveals a statistically significant effect at all time points except for the low dose at week 4. The magnitude of effect is similar to that observed for OSAHS Studies.

		Combined Armodafinil (n=118)	Placebo (n=58)	p-value
	Baseline	16.5	17.5	
Week 4	Mean Change from baseline	-3.3	-2.2	0.0282
	Median Change	-2.0	-2.0	

	from baseline			
Week 8	Mean Change from baseline	-3.2	-1.4	0.014
	Median Change from baseline	-3.0	-1.0	
Week 12	Mean Change from baseline	-4.1	-1.4	0.0002
	Median Change from baseline	-3.0	-1.0	
Final Measurement	Mean Change from baseline	-3.9	-1.9	0.0006
	Median Change from baseline	-3.0	-2.0	

- Mean baseline values and changes in the mean BFI from baseline to the endpoint for the full analysis set for placebo and combined armodafinil dose groups are presented in the table below. Note there was a statistically significant reduction in the BFI score (ANCOVA), presumably indicting a reduction in fatigue. . As noted above, the meaning of this is not clear as this metric has not been validated for this population of patients. Each individual dose was also noted to have a statistically significant reduction (data not shown). Moreover, examination of weeks 2, 8 and 12 endpoints revealed a similar statistically significant difference (not corrected for multiple comparisons) between placebo and the combined dose groups or each individual dose. There were no obvious **signs of tolerance of armodafinil's effect** on the BFI over a period of 12 weeks.

	Combined Armodafinil (n=118)	Placebo (n=58)
Mean BFI	5.6	5.7
Endpoint change from baseline	-1.4	-0.3
p value	0.0002	

- A descriptive statistical presentation of information derived from patient diaries pertinent to efficacy are presented in the table below. Percent changes for the combined armodafinil dose groups for the number of sleep, both unintentional sleep and intentional naps, were less then the placebo. This is consistent with a therapeutic effect. The effect was similar in both dose groups (data not shown). A similar number of accidents were reported between both treatment groups. Baseline caffeine use was rather high, with the armodafinil group being somewhat higher then placebo group, but little or no change was observed in both placebo and armodafinil groups during the treatment period. There were no appreciable differences in the number of episodes of cataplexy reported following treatment in both groups.

	Combined Armodafinil	Placebo

Unintentional Sleep episodes	Mean episodes at baseline	2.0	2.2
	Mean % change from baseline at endpoint	-38.7%	-10.2%
Number of daily naps	Mean number of naps at baseline	1.3	1.4
	Mean % change from baseline at endpoint	-42.5%	-22.0%
Percent of patients reporting mistakes Accidents and near misses			
	At endpoint	86%	86%
Caffeine beverage use	Mean number at baseline	16.2	13.9
	Mean number change	-1.1	0.6
Number of cataplexy attacks	Mean number	0.8	0.7
	Mean Number change	-0.1	-0.1

1.17.4.2.2.3 Secondary Endpoints Related to Cognition (CDR)

1.17.4.2.2.3.1 Key Secondary Endpoint – Quality of Episodic Memory

Improvement was observed in the Quality of Episodic memory (average of 4 tests at 0930, 1130, 1330, and 1530) in the armodafinil group as compared to the placebo group at the final endpoint evaluation. This improvement proved to be statistically significant (ANCOVA) for the combined dose and each individual dose group. These data are presented in the table below. Similar trends were observed during test periods at weeks 4, 8 and 12, but these were not always statistically significant.

Appears This Way
 On Original

Time point* Statistic	Armodafinil 250 mg/day (N=60)	Armodafinil 150 mg/day (N=58)	Armodafinil combined (N=118)	Placebo (N=58)
Baseline				
n	59	57	116	58
Mean	166.4	159.3	162.9	163.3
SD	48.31	56.81	52.55	56.12
Median	170.8	157.1	164.6	156.9
Min, max	33.7, 296.7	32.9, 272.9	32.9, 296.7	49.2, 349.6
Endpoint				
n	59	57	116	57
Mean	182.9	180.1	181.5	165.7
SD	57.49	53.22	55.21	59.80
Median	187.1	175.0	181.7	166.2
Min, max	38.8, 320.0	72.5, 289.6	38.8, 320.0	33.7, 372.1
Change from baseline to endpoint				
n	59	57	116	57
Mean	16.5	20.7	18.6	1.0
SD	46.53	34.46	40.92	29.14
Median	17.5	14.6	16.5	-5.4
Min, max	-119.2, 190.0	-46.7, 126.7	-119.2, 190.0	-66.3, 57.5
p-value	0.0168	0.0062	0.0032	—

1.17.4.2.2.3.2 Other Cognitive CDR Testing

- Speed of memory was significantly improved in the armodafinil group over placebo in the combined dose group at the final endpoint evaluation (average of 4 tests at 0930, 1130, 1330, and 1530) with a mean change from baseline observed for armodafinil of 199.7 msec versus -6.3 msec for placebo. The significant difference was also observed in the high but not low dose group. Similar results were observed for weeks 4, 8 and 12.
- Power of attention evaluation at the final endpoint measure revealed an improvement in the armodafinil group over placebo as indicated by a mean change from baseline of 41.5 msec among patients treated with armodafinil compared with an increase of 158.0 msec among patients treated with placebo. This difference was statistically significant (ANCOVA, p=0.0498) for the combined drug group but not the high dose group. Although there was a similar trend at weeks 4, 8 and 12 these were not observed to be statistically significant.
- No statistically significant difference in the continuity of attention was observed

1.17.4.3 SWSD Study (3022)

1.17.4.3.1 Demographic and Baseline Characteristics

Mean Age, sex and BMI were well matched across experimental groups. There was a slight preponderance of patients of the “white race” in placebo group, but it is doubtful this difference will affect final results.

Demographic variable Statistic, n (%)	Armodafinil 150 mg (N=123)	Placebo (N=122)	Total (N=245)	p-value
Age, years				
n	123	122	245	
Mean	38.9	40.3	39.6	0.2844 ^a
SD	10.75	10.76	10.76	
Median	37.0	41.0	41.0	
Min, max	18.0, 63.0	19.0, 62.0	18.0, 63.0	
Age group, n (%)				
18-29	31 (25)	26 (21)	57 (23)	0.2189 ^b
30-40	38 (31)	27 (22)	65 (27)	
41-55	48 (39)	59 (48)	107 (44)	
>55	6 (5)	10 (8)	16 (7)	
Sex, n (%)				
Male	66 (54)	64 (52)	130 (53)	0.8508 ^b
Female	57 (46)	58 (48)	115 (47)	
Race, n (%)				
White	74 (60)	86 (70)	160 (65)	0.3362 ^c
Black	37 (30)	26 (21)	63 (26)	
Asian	2 (2)	1 (<1)	3 (1)	
Pacific Islander	0	1 (<1)	1 (<1)	
Other ^d	10 (8)	8 (7)	18 (7)	
Race group, n (%)				
White	74 (60)	86 (70)	160 (65)	0.0895 ^b
Nonwhite	49 (40)	36 (30)	85 (35)	
BMI, kg/m²				
n	123	122	245	
Mean	29.1	30.2	29.6	0.2019 ^a
SD	6.10	7.22	6.69	
Median	27.8	29.0	28.3	
Min, max	13.3, 51.4	17.1, 57.6	13.3, 57.6	

SOURCE: Summary 15.2, Listing 5.

^a The p-value for the treatment comparison is from an analysis of variance (ANOVA) with treatment group as a factor.

^b The p-value for the treatment comparison is from a Pearson's chi-square test.

^c The p-value for the treatment comparison is from a Fisher's exact test.

^d Other=Hispanic (15 patients); black/Asian (1 patient), Bangladeshi (1 patient); Caucasian/African American (1 patient).

BMI=body mass index; min=minimum; max=maximum.

Baseline characteristics are presented in the table below. Job types, occupation and shift worker type were generally similar across treatment groups. Most subjects worked as permanent night

workers (evening shift only). CGI-S scores were similar across experimental groups. A majority of patients were rated as moderately ill with the remainder receiving higher ratings. **The degree of "illness" observed in these trials was** similar to that in the trials of the other disorders. Sleepiness, as measured by the MSLT, was comparable across treatment groups (2.3 minutes for armodafinil and 2.4 minutes for placebo).

Appears This Way
On Original

Variable Category	Armodafinil		Total (N=245)	p-value
	150 mg (N=123)	Placebo (N=122)		
CGI-S				
Normal-not at all ill	0	0	0	0.8988 ^a
Borderline ill	0	0	0	
Slightly ill	0	0	0	
Moderately ill	69 (56)	69 (57)	138 (56)	
Markedly ill	42 (34)	44 (36)	86 (35)	
Severely ill	11 (9)	8 (7)	19 (8)	
Among the most extremely ill	1 (<1)	1 (<1)	2 (<1)	
CGI-S group, n (%)				
Moderately ill	69 (56)	69 (57)	138 (56)	0.9422 ^b
Markedly, severely, or extremely ill	54 (44)	53 (43)	107 (44)	
Shift worker type				
Permanent	107 (87)	105 (86)	212 (87)	
Rotating	16 (13)	17 (14)	33 (13)	
Shift worker occupation				
Mining	1 (<1)	0	1 (<1)	
Utilities	1 (<1)	3 (2)	4 (2)	
Construction	1 (<1)	0	1 (<1)	
Manufacturing	10 (8)	5 (4)	15 (6)	
Wholesale trade	0	1 (<1)	1 (<1)	
Retail trade	3 (2)	4 (3)	7 (3)	
Information	3 (2)	1 (<1)	4 (2)	
Transportation and warehousing	11 (9)	9 (7)	20 (8)	
Postal service	4 (3)	2 (2)	6 (2)	
Agriculture, forestry and fishing	1 (<1)	0	1 (<1)	
Finance and insurance	1 (<1)	1 (<1)	2 (<1)	
Professional, scientific and legal	7 (6)	10 (8)	17 (7)	
Management of companies and offices	0	2 (2)	2 (<1)	
Administrative and support services	8 (7)	8 (7)	16 (7)	
Health care and social assistance	51 (41)	47 (39)	98 (40)	
Arts, entertainment and recreation	2 (2)	1 (<1)	3 (1)	
Accommodation and food services	4 (3)	7 (6)	11 (4)	
Other services (except public administration)	15 (12)	19 (16)	34 (14)	
Public administration	0	2 (2)	2 (<1)	

SOURCE: Summary 15.3; Listing 6 and Listing 12.

^a The p-value for the treatment comparison is from a Fisher's exact test.

^b The p-value for the treatment comparison is from a Pearson's chi-square test.

CGI-S=Clinical Global Impression of Severity.

1.17.4.3.2 Efficacy Results

1.17.4.3.2.1 Primary Endpoints

1.17.4.3.2.1.1 MSLT

Analysis of sleep latency using the MSLT resulted in very few flawed sleep sessions for reasons previously described. The small number of flawed sessions is unlikely to corrupt the final efficacy conclusions. Nonetheless, the table below presents post-hoc analyses along with the original NDA analysis. The worst case scenario was not requested by statistics, and that analysis was not performed, presumably because of the small number of flawed sessions. Since there was evidence of treatment covariate interaction the covariate was dropped from the ANCOVA model and an ANOVA was used. The table below presents the change from baseline to the final endpoint evaluation of both experimental groups for the original analysis as well as the post-hoc analyses. The magnitude of effect is similar amongst all analyses with a prolongation in sleep latency of 2.3 to 3.3 for the armodafinil group over the placebo group. All analyses exhibited a similar p value ($p < 0.001$).

Appears This Way
On Original

Method *#	Armodafinil 150 mg	Placebo	Treatment difference 95% CI
	N Mean (SD) p-value	N Mean (SD)	
Analysis 1	112 3.0 (4.38) <0.0001	104 0.4 (2.86)	1.57, 3.56
Analysis 2	112 3.0 (4.35) <0.0001	104 -0.3 (2.89)	1.69, 3.68
Analysis 3	112 2.9 (4.30) <0.0001	104 0.3 (2.87)	1.54, 3.51
Analysis 4	112 2.8 (4.18) <0.0001	104 0.3 (2.87)	1.52, 3.45
Analysis 5	107 2.7 (4.12) <0.0001	103 0.4 (2.87)	1.37, 3.30
Analysis 6	112 2.9 (4.35) <0.0001	104 0.3 (2.87)	1.58, 3.57
Original NDA	112 3.1 (4.46) <0.0001	104 0.4 (2.87)	1.67, 3.69

Because of the model residuals showed evidence of nonnormality, a Wilcoxon rank-sum test was performed using the original NDA dataset to assess the robustness of the inference based on the ANOVA model fit. The Sponsor noted that inferences from both parametric (ANOVA) and nonparametric (Wilcoxon rank-sum test) tests were consistent in that they both demonstrated significant p-values (<0.05). The Sponsor notes that in accordance with the statistical analysis plan, all efficacy results and the discussion in this report are based on the ANOVA procedures.

1.17.4.3.2.1.2 CGI-C

The endpoint of the percent of patients with at least minimal improvement is presented in the table below. More patients reached these criteria in the armodafinil treatment. This difference was found to be statistically significant (Cochran-Mantel-Haenzel chi-square).

Variable, n (%)	Armodafinil		p-value
	150 mg (N=112)	Placebo (N=104)	
At least minimal improvement ^a	89 (79)	61 (59)	0.0010
No improvement	23 (21)	43 (41)	

Frequencies of various CGI-C scores are presented below. Like the other disorders studied, most of the differences between the experimental groups lie in the greater percent of patients who are judged as much improved and very much improved in the armodafinil group.

Rating, n (%)	Armodafinil	
	150 mg (N=112)	Placebo (N=104)
Very much improved	25 (22)	13 (13)
Much improved	39 (35)	24 (23)
Minimally improved	25 (22)	24 (23)
No change	20 (18)	38 (37)
Minimally worse	0	5 (5)
Much worse	3 (3)	0
Very much worse	0	0

1.17.4.3.2.2 Secondary Endpoints Related to Sleep and Fatigue

- The table below presents the change from baseline in sleep latency measured as determined by the MSLT (average of 4 naps at 0200, 0400, 0600, and 0800) assessed at weeks 4, 8, and 12 for armodafinil versus placebo. As is apparent, statistically significant changes were observed throughout the experimental period (ANOVA, not corrected for multiple comparisons). An effect was seen at the first test period and there is no obvious tolerance throughout the period

	Week 4		Week 8		Week 12	
	Armodafinil (n= 112)	Placebo (n= 104)	Armodafinil (n=101)	Placebo (n=94)	Armodafinil (n=87)	Placebo (n=83)
Mean Change from baseline (min)	3.1	0.6	3.4	0.1	3.4	0.5
p value	<0.0001		<0.0001		<0.0001	

- CGI-C was evaluated at each test period. The data are presented below. Armodafinil effect was significant at all test periods (Cochran-Mantel-Haenzel chi-square test, not corrected for multiple comparisons). **Armodafinil's effect appears to be greatest at week 8**, with weeks 4 and 12 exhibiting effects of similar magnitude. These differences in effect are unlikely to be clinically significant.

Time point Variable, n (%)	Armodafinil		p-value
	150 mg (N=112)	Placebo (N=104)	
Week 4	110 (100)	100 (100)	
At least minimal improvement ^a	89 (81)	59 (59)	0.0005
No improvement	21 (19)	41 (41)	
Week 8	99 (100)	93 (100)	
At least minimal improvement ^a	77 (78)	45 (48)	<0.0001
No improvement	22 (22)	48 (52)	
Week 12	96 (100)	89 (100)	
At least minimal improvement ^a	75 (78)	50 (56)	0.0014
No improvement	21 (22)	39 (44)	

- The table below presents the KSS change from baseline to various treatment periods throughout the study. Armodafinil produced a statistically significant (ANOVA) greater reduction in the KSS score over placebo at all time points indicates a therapeutic reduction in sleepiness. There was a slight decrement of magnitude in response over time, but it is uncertain if this reduction would be significant.

	Armodafinil (n=112)	Placebo (n=104)	p-value
Baseline	7.4	7.3	
Mean Change from baseline at 4 Weeks	-1.8	0.8	<0.0001
Mean Change from baseline at 8 weeks	-1.7	-0.7	<0.0001
Mean Change from baseline at 12 weeks	-1.9	-1.1	0.0034
Mean Change from baseline at endpoint	-1.8	-1.0	0.0008

- Results for the BFI at final endpoint evaluation are presented in the table below. Although there was a greater reduction in the BFI in the armodafinil group this reduction was small and not statistically significant.

	Armodafinil (n=112)	Placebo (n=104)
Mean BFI	7.5	7.4
Endpoint change from baseline	-1.1	-0.8
p value	0.3475	

- Information from patients diaries are presented in the table below. All patients appeared to experience a decrease in unintentional episodes of sleep and intentional naps during the experimental period. This reduction, however, was most apparent in patients on armodafinil suggesting a therapeutic effect. There was no difference between both experimental groups in the percent of patients experiencing accidents or near misses at work or during the commute home. The armodafinil group had a slightly lower use of caffeine beverages but also experienced the largest reduction in caffeine beverage use during the experimental phase.

		Armodafinil	Placebo
Unintentional Sleep episodes	Mean episodes at baseline	1.2	1.1
	Mean % change from baseline at endpoint	-71.8%	-42.2%
Number of daily naps	Mean number of naps at baseline	0.7	0.6
	Mean % change from baseline at endpoint	-35.8%	-13.2%
Percent of patients reporting mistakes Accidents and near misses	During Night shift	65%	65%
	On the Commute Home	50%	50%
Caffeine beverage use	Mean number at baseline	1.3	1.8
	Mean number change	-0.4	0.0

1.17.4.3.2.3 Secondary Endpoints related to Cognition (CDR)

1.17.4.3.2.3.1 Key Secondary Endpoint: Quality of Episodic Memory

The table below presents the difference between the 2 treatment groups at baseline and endpoint evaluations for the key secondary endpoint, the quality of episodic memory that was calculated as the average of four tests between times 0230 to 0830. There was an improvement in results for drug and a decrement for placebo. These results were found to be statistically significant

(ANOVA, $p < 0.0001$) in favor of the armodafinil group. Testing at weeks 4, 8 and 12 demonstrated a similar magnitude and statistically significant result (data not shown).

Time point^a Statistic	Armodafinil 150 mg (N=112)	Placebo (N=104)	p-value
Baseline			
n	109	100	
Mean	141.5	138.3	
SD	49.12	50.23	
Median	150.4	139.0	
Min, max	-40.7, 224.2	-44.6, 311.3	
Endpoint			
n	110	101	
Mean	159.9	135.1	
SD	53.51	59.84	
Median	163.1	135.8	
Min, max	-33.3, 302.9	-17.9, 352.5	
Change from baseline to endpoint			
n	109	100	
Mean	18.4	-3.3	<0.0001
SD	38.34	39.62	
Median	17.1	-6.0	
Min, max	-115.0, 118.7	-136.7, 101.2	

1.17.4.3.2.3.2 Other Cognitive CDR Testing

- While Speed of Memory testing in the CDR tool numerically suggested improvement (116.8 msec difference between experimental groups) in the armodafinil group as compared to the placebo group at the final endpoint measure, the result was not statistically significant (ANOVA, data not shown). A similar statistically insignificant trend for armodafinil for improvement with armodafinil was observed at week 4, 8 and 12 evaluations.
- The power of attention (a function of speed of attention) from the CDR system (average of 4 tests at 0230, 0430, 0630, and 0830) was observed to be statically significantly improved (ANOVA, $p=0.0011$) in the armodafinil group as compared to the placebo group with a with a mean decrease of 88.3 msec in armodafinil-treated patients and mean increase of 88.1 msec in placebo-treated patients at the final endpoint evaluation. Similarly statistically significant changes of similar magnitudes were observed for each individual testing period (weeks 4, 8 and 12).

- The continuity of attention from the CDR system (average of 4 tests at 0230, 0430, 0630, and 0830) endpoint was observed to be statistically significantly (ANOVA, $p=0.0005$) improved at the final endpoint evaluation. The mean change from baseline to endpoint in continuity of increased by 2.9 in the armodafinil-treated patients and 0.2 in the placebo-treated patients. Similar magnitude of effect and statistical significance was observed at week 8 and 12. There was a trend in the right direction of effect but this was not observed to be statistically significant at week 4.

1.17.5 Clinical Microbiology

Does not Apply.

1.17.6 Efficacy Conclusions

The current series of studies on the therapeutic benefit of armodafinil in the treatment of sleepiness associated with the three dyssomnias of narcolepsy, OSAHS and SWSD was complicated by a flaw in the database.

As in the past, this division had requested that the Sponsor finds both of two primary endpoints, one subjective (the CGI-C) and one subjective (sleep latency), significant to be awarded a therapeutic claim. The subjective endpoint, patients with at least a minimal improvement in the CGI-C was observed to be significant every study for every dose examined. There was little difference between the therapeutic effects of either the high or low dose. As noted above these effects are principally attributed to patents with much to very much improved scores. A summary table presented these data are presented below.

Disorder	NUVIGIL 150 mg*	NUVIGIL 250 mg*	Placebo
OSAHS I	71%	74%	37%
OSAHS II	71%	-	53%
Narcolepsy	69%	73%	33%
SWSD	79%	-----	59%

*Significantly different than placebo for all trials ($p<0.01$)

These effects are very similar to those observed for the racemate Provigil which is presented below (data from label).

Disorder	PROVIGIL 200 mg [*]	PROVIGIL 400 mg [*]	Placebo
Narcolepsy I	64%	72%	37%
Narcolepsy II	58%	60%	38%
OSAHS	61%	68%	37%
SWSD	74%	—	36%

^{*}Significantly different than placebo for all trials (p<0.01)

The problem arose when it was discovered that a number of flaws existed in the sleep latency database. The incidence of this event in the SWSD study was relatively rare (0.2 to 9.8%) and probably does not significantly affect the final conclusion for statistically significant prolongation in latency and, therefore, affect a therapeutic claim. The incidence of this effect the OSAHS and narcolepsy MWT sessions were, however, significant with, 6.2% to 15.3% of sessions being flawed. These flawed sessions resulted from truncation of such sessions by local readers because patients were scored as reaching the protocol driven criteria. The disparity to central readers, however, resulted in these patients being scored as not having fallen asleep and therefore were given a maximal latency score. This reviewer believes that this disparity should have been brought to the attention of the agency. Nonetheless, this reviewer believes that the data still indicates a statistically significant improvement in both doses of the armodafinil. For the following reasons:

- The flawed sessions appeared to be randomly distributed amongst the various dosage and placebo groups.
- Every one of the non-worst case scenario post-hoc analyses requested by the division demonstrate a statistically significant therapeutic difference of similar magnitude to that observed for the original NDA analysis.
- The worst case scenario, which restricted itself to the use of the first 4 sleep session (7.2), proved to demonstrate a statistically significant therapeutic effect in all but one case. The worst case scenario that utilized latter sleep sessions still exhibited a trend in therapy but exhibited only one value that was statistically significant. This would not be unexpected the inclusion of later testing values **would dilute armodafinil's effect because** of declining blood levels.

The conservative worst case scenario (8.1 and 8.2) failed to demonstrate an effect. This reviewer believes that this test may have been biased to an extreme.

It is noteworthy that in the initial review the statistician felt that there was no evidence of **efficacy. They noted that they "would feel more confident to support the efficacy claim if at least some of the worst case analyses are statistically significant."** In discussions this reviewer had with the statistician it became apparent that the significance of the worst case scenario analyses was missed; they had assumed no significance. They subsequently agreed with this significance and a therapeutic claim. They have added an addendum to their review confirming this.

In conclusion this reviewer believes that analysis of sleep latency in each pivotal study indicated a therapeutic effect. The table below summarizes this effect, using the original dataset analysis.

Disorder	Measure	NUVIGIL 150 mg *		NUVIGIL 250 mg*		Placebo	
		Baseline	Change from Baseline	Baseline	Change from Baseline	Baseline	Change from Baseline
OSAHS I	MWT	21.5	1.7	23.3	2.2	23.2	-1.7
OSAHS II	MWT	23.7	2.3	-	-	23.3	-1.3
Narcolepsy	MWT	12.1	1.3	9.5	2.6	12.5	-1.9
SWSD	MSLT	2.3	3.1	-	-	2.4	.4

*Significantly different than placebo for all trials (p<0.01)

The magnitude of effect (drug compared to placebo) is rather similar to that observed for Provigil. The table below presents the Provigil⁴ data (from label).

Disorder	Measure	PROVIGIL 200 mg *		PROVIGIL 400 mg *		Placebo	
		Baseline	Change from baseline	Baseline	Change from baseline	Baseline	Change from baseline
Narcolepsy I	MWT	5.8	2.3	6.6	2.3	5.8	-0.7
Narcolepsy II	MWT	6.1	2.2	5.9	2.0	6.0	-0.7
OSAHS	MWT	13.1	1.6	13.6	1.5	13.8	-1.1
SWSD	MSLT	2.1	1.7	-	-	2.0	0.3

*Significantly different than placebo for all trials (p<0.01 for all trials but SWSD, which was p<0.05)

Results from secondary endpoint sleepiness rating scales, ESS and KSS, support the conclusion of armodafinil therapeutic effect. The conclusion of efficacy is also supported by the descriptive **observation from patient's diaries that indicate** reduced unintentional episodes of sleep and reduced intentional napping. The information on **"accidents and near misses," in these diaries** did not always show improvement with armodafinil treatment. Some studies did show some improvement in this measure with armodafinil group.

Evaluation of endpoints during the complete 12 week experimental period does not indicate a tolerance to the wakefulness promoting effect of armodafinil. The effects of the drug was apparent at the first time of measurement (4 weeks).

Two "cognitive" key secondary endpoint where identified by the Sponsor. Both where derived from CDR testing. The first, **Power of Attention"** was identified as a key secondary endpoint for the OSAHS study 3025. No statistically significant improvement in this endpoint, in the

4 A similar problem with datasets may be present with the original Provigil data. The Sponsor has been queried about this and further follow-up may be necessary. Nonetheless, the present examination of data would indicate that this problem may not negatively affect final conclusions.