FDA Executive Summary

Prepared for the March 21, 2019 Meeting of the
Neurological Devices Panel
Gaithersburg Hilton; Gaithersburg, MD

De Novo DEN160053
Neuronix, Ltd.
neuroAD™ Therapy System

Division of Neurological and Physical Medicine Devices
Office of Device Evaluation
Center for Devices and Radiological Health
Food and Drug Administration

Table of Contents

List of Figures

List of Tables

List of Panel Questions with Tracing in Document

As discussed throughout this document, the Panel will be asked to discuss and make recommendations on the following:

1. Whether the U.S. pivotal study demonstrates a clinically meaningful benefit for the neuroAD as an adjunctive therapy **[page 63]**

2. When the neuroAD is used as an adjunctive therapy, the Panel will be asked to discuss and make recommendations on what minimum amount of improvement in ADAS-Cog alone is clinically meaningful, as well as the minimum amount of clinically meaningful improvement in the CGIC. **[page 10]**

3. Whether the ADAS-Cog≤30 population is a clinically plausible subset **[page 77]**

4. Whether the post-hoc identification of the ADAS-Cog≤30 population at a later time point when no treatment is given is an adequate analysis of the pivotal study data, in concert with the supplemental data provided, to demonstrate probable benefit. **[page 77]**

5. Whether the probable benefits to health outweigh the probable risks **[page 82 and 91]**

6. Whether the proposed indications for use is supported by the data collected in the clinical studies. We ask that this includes a consideration regarding selecting potential patients using the ADAS-Cog **[page 8]**

Introduction

This is FDA's Executive Summary regarding the de novo application from Neuronix, Ltd. for the neuroAD Therapy System (hereafter referred to as "neuroAD" or "the device"). The neuroAD is proposed as an adjunctive treatment of mild to moderate dementia of the Alzheimer's type in patients with a baseline ADAS-Cog score ≤30.

The neuroAD provides transcranial magnetic stimulation (TMS) delivered concurrently with cognitive training displayed on a computer touch screen. In the neuroAD, the TMS and the cognitive training is intended to target the same brain region to deliver a net therapeutic effect. Although the mechanism of action of TMS has not been established, Neuronix provides a hypothesis that the TMS component of the therapy may modulate synaptic activity, potentially enhancing the therapeutic effect of cognitive training. Treatment sessions for the neuroAD are conducted in the clinic, where the patient must go to the clinic to receive therapy five days per week for six weeks. Each daily treatment session is expected to last around one hour.

This document presents background on the regulatory history of the subject submission, a description of the device, the regulatory standards, the results of the primary clinical data set from the 106-patient pivotal study, and an analysis of the benefit/risk profile of the device. The Appendix of this document includes information on the pre-clinical studies, presents the results from additional post-hoc analyses using supplemental clinical data sets, and provides a summary of the available stakeholder input from a patient/caregiver survey and a physician survey provided by Neuronix.

The focus of this panel is the clinical data submitted for the neuroAD in order to market the device in the United States. Neuronix has provided clinical data from a pivotal study that was analyzed per the pre-specified statistical analysis plan as well as through additional post-hoc analyses. The sponsor has also provided supplemental investigations that include pilot studies, independently run clinical studies, and commercial clinic cases performed outside the United States (OUS) and provided pooled analyses of these sources. The FDA will be seeking panel input on the benefit-risk profile of the device given the clinical data that has been presented. This includes interpretation of: 1) the pre-specified primary endpoint of the pivotal study (ADAS-Cog[1] change from baseline at 7-weeks) that demonstrated non-statistically significant improvements in the sham group compared to the neuroAD treatment group (sham vs. treatment, -1.38 points vs. 0.07 points, p=0.09); 2) the pre-specified secondary endpoint assessments of the pivotal study (ADAS-Cog change from baseline at 12 weeks after the treatment ended at 6 weeks (sham vs. treatment, -0.61 points vs. -1.03 points, p=0.64) and CGI-C[2] at 7 weeks (sham vs. treatment, 4.06 points vs. 4.04 points, p=0.96) and 12 weeks (sham vs. treatment, 4.19 points vs. 3.84 points, p=0.12); 3) post-hoc analysis of the pivotal study (see Pivotal Clinical Study – Post-Hoc Analysis of Baseline ADAS-Cog≤30 Subgroup); and 4) analyses of supplemental datasets (see APPENDICES II through IV).

---

[1] On the 70-point Alzheimer's Disease Assessment Scale – cognitive subscale (ADAS-Cog) a higher score represents greater impairment in memory and other aspects of cognition; therefore, negative numbers indicate improvement on the scale.

[2] On the 7-point Clinical Global Impression - Change scale (CGI-C) a 1 indicates very much improved; 2-much improved; 3-minimally improved; 4-no change; 5-minimally worsened; 6-much worsened; or 7-very much worsened.

The FDA will also be seeking input on the high level of uncertainty in the methodology and conclusions drawn from the pivotal study subgroup post-hoc analyses and the post-hoc analyses using pooled clinical data. Given the primary results of the pivotal study did not demonstrate device effectiveness, the post-hoc analyses form the basis for the company's conclusion of device effectiveness in the intended population.

The FDA understands the current unmet need for patients suffering from Alzheimer's disease. The FDA Commissioner, Dr. Scott Gottlieb, has listed Alzheimer's disease as an area in urgent need of drug regulatory reform. Likewise, the Center for Devices and Radiological Health (CDRH), and more specifically, the Division of Neurological and Physical Medicine Devices (DNPMD) in the Office of Device Evaluation (ODE) recognizes the unmet need in the Alzheimer's disease device landscape. DNPMD also must balance the urgent need for therapies with the demonstration of safety and effectiveness.

In reviewing the totality of the evidence, there is significant uncertainty in interpreting the primary results of the pivotal study and the supplemental post-hoc analyses in assessing the proposed marketing indication. FDA would appreciate panel input on issues regarding device safety, effectiveness, clinically meaningful benefit for an adjunctive Alzheimer's disease therapy, and the overall benefit-risk ratio of the neuroAD.

Proposed Indications for Use

The proposed indications for use (IFU) for the neuroAD™ device is:

> The neuroAD™ Therapy System is intended for neuro-stimulation concurrently combined with cognitive training. neuroAD™ Therapy System is indicated for the treatment of mild to moderate dementia of the Alzheimer's type in patients with a baseline ADAS-Cog score up to 30. neuroAD™ Therapy System may be used in conjunction with other pharmacological and non-pharmacological therapies.

Panel Question: The Panel will be asked to discuss and make recommendations on whether the proposed indications for use is supported by the data collected in the clinical studies. We ask that this includes a consideration regarding selecting potential patients using the ADAS-Cog.

Clinical Context

Alzheimer's Dementia (AD) is a progressive neurodegenerative disorder that impairs memory, thinking, language and behavior. It is characterized by cognitive impairment, progressive neurodegeneration and formation of amyloid-b (Ab)-containing plaques and neurofibrillary tangles composed of hyperphosphorylated tau. Disruption of hippocampal circuitry leads to the inability to consolidate immediate and short-term memory into long-term traces. Temporal lobe damage and disruption of connections with the basal forebrain, cingulate cortex, frontal cortex, and other forebrain structures affected by AD contribute to the marked cognitive decline in patients with AD. In AD, the brain shows extensive neuronal loss, impaired functioning of synaptic connections, and damage to important neurotransmitter systems that participate in functions such as memory. The most common and often earliest clinical symptom is selective memory impairment. Declarative episodic memory, which depends on the hippocampus and medial temporal lobe, is usually affected early on in AD. The impairment of executive function, judgement, and problem solving are additional clinical manifestations and can appear in the early stages. Family and coworkers due to difficulty with abstract reasoning and an inability to complete tasks primarily report these symptoms. The patient, due to reduced insight into their deficits (anosognosia), may not recognize these or other symptoms. The more impaired an individual's insight is

the more likely they will become agitated, disinhibited or exhibit issues which impact safety (such as driving). Behavioral and neuropsychiatric symptoms may manifest as well in AD. Neuropsychiatric symptoms may appear with subtlety as apathy, disengagement, or irritability. Behavioral symptoms including aggression, wandering and various psychotic manifestations (hallucinations, delusions, and misidentification/misperception) may occur usually during the mid to later stages of AD progression. These symptoms are common in AD, and lead to greater functional impairment.

In this way, Alzheimer's Dementia is a syndrome that not only has direct effects upon the patient, it affects the patient's entire family and social support structure. The debilitating deficits in AD may not take its toll initially on the patient's physical capabilities but does gradually become debilitating in their ability to engage with others and their environment. It takes away their ability to perform the daily activities of life. The current available treatments do not cure or slow the progression of neurodegeneration.

Alzheimer's is the most common dementia in the United States (US) and worldwide. The most common risk factor for Alzheimer's is age, particularly in those over age 65. The National Center for Health Statistics (NCHS) projects that by 2030, 1 in 5 Americans, or 72.7 million, will be aged 65 years or older, and by 2050 the number of adults aged 65 years or older in the United States is projected to reach 83.7 million. An estimated 5.7 million Americans of all ages are living with Alzheimer's dementia in 2018. This number includes an estimated 5.5 million people aged 65 and older and approximately 200,000 individuals under age 65 who have younger-onset Alzheimer's. By 2050, the total number is projected to rise to 14 million people.

Clinical Assessment Scales for Alzheimer's Dementia

The Alzheimer's Disease Assessment Scale - cognitive subscale (ADAS-Cog)

The Alzheimer's Disease Assessment Scale (ADAS) intends to measure the severity of the most characteristic cognitive symptoms of Alzheimer's disease (AD). Its subscale ADAS-Cog is the most popular cognitive testing instrument used in clinical trials of nootropics. It consists of 11 tasks measuring the disturbances of memory, language, praxis, attention and other cognitive abilities often referred to as the core symptoms of AD. The total score ranges 0-70 points and measures the number of mistakes counted in the test. A higher score represents greater impairment in memory and other aspects of cognition.

The test administrator adds up points for the errors in each task of the ADAS-Cog for a total score. The greater the dysfunction, the greater the score.

In assessing a patient with the ADAS-Cog, the scale benefits from the entirety of data gathered on the patient in multiple areas to provide a picture of the patient's level of function in different areas of cognition. A person with a score of up to five on the ADAS-Cog could be considered "normal". With considerations for age, education and other neuropsychological testing norms, a score of >=18 is considered impaired.

For example, in the ADAS-Cog section, testing word recall, the patient is given three trials to learn a list "high frequency, high imagery" nouns. All 10 words are printed on white cards. The score for this section is the mean number of words NOT recalled on three trials with a maximum score of ten. A mean change of remembering two additional words on this section results in a 2 point improvement on the overall ADAS-Cog score. When taking a mean change of one to two points on the overall score for a single patient over time, the clinical change in the patient may not be exhibited in their overall or global function. As with other standard neuropsychological or neurocognitive testing there can be issues related to introducing external variables or confounders that affect the results.

Practice effects (PEs) which are gains due to prior exposure to tests or assessments can be the result of a patient taking a neuropsychological or cognitive assessment multiple times, particularly in a shortened time interval. Although the patient may not "memorize" the "answers", the patient can become familiar with the test and know how to approach the tasks more efficiently. This could result in the masking, inflating or skewing of the results of the intervention being studied. It also appears that the test-retest reliability of ADAS-Cog for naïve raters may differ from that of experienced raters at well-trained sites. The potential PEs that can occur with serial testing in fairly short intervals (weeks apart) and the test-retest reliability should be considered when interpreting trial results.

The Clinical Global Impression - Change scale (CGI-C)

The Clinical Global Impression - Change scale (CGI-C) is a 7-point scale that requires the clinician to rate the severity of the patient's illness at the time of assessment, relative to the clinician's past diagnosis with the patient. Considering total clinical experience, a patient was assessed on severity of mental illness at the time of rating. Assessment was performed using a structured interview with both the patient and the caregiver. The purpose of the baseline measurement of CGI-C was to set the reference for future comparison. On the scale of 1-7: 1-very much improved; 2-much improved; 3-minimally improved; 4-no change; 5-minimally worsened; 6-much worsened; or 7-very much worsened.

Panel Question: When the neuroAD is used as an adjunctive therapy, the Panel will be asked to discuss and make recommendations on what minimum amount of improvement in ADAS-Cog alone is clinically meaningful, as well as the minimum amount of clinically meaningful improvement in the CGIC.

Below is a brief tabular summary of the clinical assessment scales used in the clinical trial. This table does not represent all of the available scales used in the assessment of AD patients.

Table 1. Clinical Assessment Scales Used to Evaluate neuroAD in Clinical Trials

Domain/Measure	Scale Name and Decription
Cognition: Memory, orientation, language, praxis, etc.	**Mini-Mental State Exam (MMSE)** **Description**: 30-pt. scale (higher scores better). Clinician administered patient evaluation **Type/Use**: Commonly used in general clinical practice. Usually 6-10 minutes to administer. Mostly used for eligibility screening and dementia staging in clinical trials. Assesses a wide range of domains, including attention, language, memory, orientation, and visuospatial proficiency.
	Alzheimer's Disease Assessment Scale-cognitive subscale (ADAS- cog) **Description**: 70-pt. scale (scores >=18 indicating cognitive impairment; the higher the score, the greater the impairment). Clinician administered patient evaluation. Measures cognitive domains including memory, language and praxis. **Type/Use**: Standard cognitive outcome measure in mild-moderate AD in clinical research.
Global Change: Summary outcome assessment from baseline to endpoint	**Clinical Global Impression of Change (CGI-C)** Also referred to as the ADCS-CGIC (Alzheimer's Disease Cooperative Study) **Description**: 7-point scale [1=very much improved; 4=no change; 7=very much worse]

	Clinician rated, based on patient +/- informant interview. It requires the assessor to consider a number of cognitive, functional, and behavioral areas prior to providing an overall "global" assessment of clinical change using a worksheet that comprehensively lists relevant symptoms potentially useful in judging clinically meaningful change and allows for notes for future reference- it takes approximately 20 minutes per interview. **Type/Use**: Outcome measure used in AD clinical trials. A systematic method for assessing clinically significant change focusing on clinicians' observations of change in the patient's cognitive, functional, and behavioral performance since the beginning of a trial.
	Clinical Dementia Rating (CDR) Also referred to as the Washington University Clinical Dementia Rating **Description**: CDR scale is a five-point scale describing subjects without dementia (0) and with questionable (0.5), mild (1), moderate (2), and severe (3) dementia The clinical protocol incorporates semi-structured interviews with the patient and informant to obtain information necessary to rate the subject's cognitive performance in six domains: memory, orientation, judgment and problem solving, community affairs, home and hobbies, and personal care. Each category is scored as independently as possible according to descriptions of impairment found in the CDR table, using the same five-point scale of impairment (0 through 3). These scores are then combined according to established rules to determine the overall CDR. **Type/Use**: A global scale developed to clinically denote the presence of AD and stage its severity. Not well suited for brief screening but used in clinical and research settings. The protocol takes about 90 minutes to administer.

Current Treatment Options for Alzheimer's Disease

The focus for AD treatment and product development has been in been maintaining mental function, managing behavioral symptoms and altering the rate of disease progression.

Currently, there are pharmacological treatments that may ameliorate or lessen some of the symptoms of AD for what appears to be a limited amount of time in mild to moderate stages of AD. Thus far, there is not a cure or disease modifying (treatment that slows the course of the illness) available.

Cholinesterase inhibitors (AChEIs) such as galantamine, rivastigmine, and donezepil are cleared for treatment of mild AD. Memantine, a N-Methyl-D-Aspartate (NMDA) receptor antagonist, is cleared for treatment in patients with moderate to severe AD. Memantine addresses dysfunction in glutamatergic transmission, while the AChEIs serve to increase pathologically lowered levels of the neurotransmitter acetylcholine. Memantine is used alone or often in combination with a cholinesterase inhibitor. These agents are considered symptomatic therapies and are not thought to be neuroprotective or to alter the underlying disease trajectory.

Table 2. FDA-approved drugs for the treatment of AD

Drug Name	Brand Name	Approved Stage	Approval Year
Donepezil	Aricept	All stages	1996
Galantamine	Razadyne	Mile to moderate	2001
Memantine	Namenda	Moderate to severe	2003
Rivistagimine	Exelon	All stages	2000
Donepezil and Memantine	Namzaric	Moderate to severe	2014

These pharmacotherapeutics do have drawbacks. The side effects of taking them are intolerable or difficult to manage in some patients. The side effects include dizziness, headache, constipation, confusion, nausea, vomiting, diarrhea, muscle cramps, indigestion, fatigue, and weight loss.

FDA Approved pharmacological treatments related to AD

Three acetylcholinesterase inhibitors are currently marketed in this country for the treatment of Alzheimer's Disease: donepezil, rivastigmine, and galantamine. Each was initially FDA approved for the treatment of mild to moderate dementia of the Alzheimer's type.

Medical Devices and AD

Currently there no approved devices for the treatment of Alzheimer's Dementia. The need is great, however the safety, benefit, and the degrees of uncertainty must be considered.

Regulatory Background

The regulatory submission that is the topic of discussion at this meeting is a de novo premarket submission. Other TMS devices that were first of a kind for major depressive disorder (MDD) and obsessive-compulsive disorder (OCD) were cleared according to the same de novo pathway.

Per the statute, a device is eligible for classification via the De Novo pathway if there is no predicate device (a legally marketed device that is not subject to Premarket Approval (PMA)) and if the device under consideration presents a low- to moderate-risk profile and general or general and special controls would provide reasonable assurance of the safety and effectiveness of the device. FDA determined that the neuroAD device met the criteria to be considered. If granted, the resulting device regulation places the device type in class I (general controls) or class II (general and special controls)

The data which CDRH considers for review is identified as valid scientific evidence. Per 21 CFR 860.7(c)(2), "valid scientific evidence is evidence from well-controlled investigations, partially controlled studies, studies and objective trials without matched controls, well-documented case histories conducted by qualified experts, and reports of significant human experience with a marketed device, from which it can fairly and responsibly be concluded by qualified experts that there is reasonable assurance of the safety and effectiveness of a device under its conditions of use." De novo applications must adhere to this standard.

In order to be granted, the evidence in the submission must demonstrate a reasonable assurance of safety and effectiveness as defined in 21 CFR 860.7(d)(1) and (e)(1), respectively. Summarized, the evidence must show that when using the device properly, the probable benefits to health outweigh any probable risks and there is an absence of unreasonable risk (safety), and that there are clinically significant results in a significant portion of the target population (effectiveness).

The FDA Guidance Document, "Factors to Consider When Making Benefit-Risk Determinations in Medical Device Premarket Approval and De Novo Classifications" (U.S. Food and Drug Administration, 2016b) is applicable to this submission. More information regarding the FDA's views regarding the neuroAD Benefit-Risk profile is included in section Benefit-Risk Assessment

In addition, since the FDA Modernization Act of 1997 (FDAMA), Congress has directed FDA to take a least burdensome approach to medical device premarket evaluation in a manner that eliminates unnecessary burdens that may delay marketing of beneficial new products, while maintaining the statutory requirements for clearance and approval.

Device Description

The neuroAD™ Therapy System (model NICE V.3) delivers treatment in the form of transcranial magnetic stimulation (TMS) applied to the head concurrently with cognitive training exercises presented on a computer screen.

The overall neuroAD system is composed of two sub-units:

- Base Unit – Contains and controls the TMS and the Cognitive Training modules. The Cognitive Training module is presented to the patient on a touch-screen.

- Navigation Unit – Directs the operator on positioning the TMS administrating coil on the spatially discrete regions of the cerebral cortex (i.e., treatment region). The treatment region is marked in advance on an individual MRI scan.

Each unit has several components that are represented in Figure 1 in green (Navigation Unit) or purple (Base Unit) labeling. Critical components are also described in greater detail below.

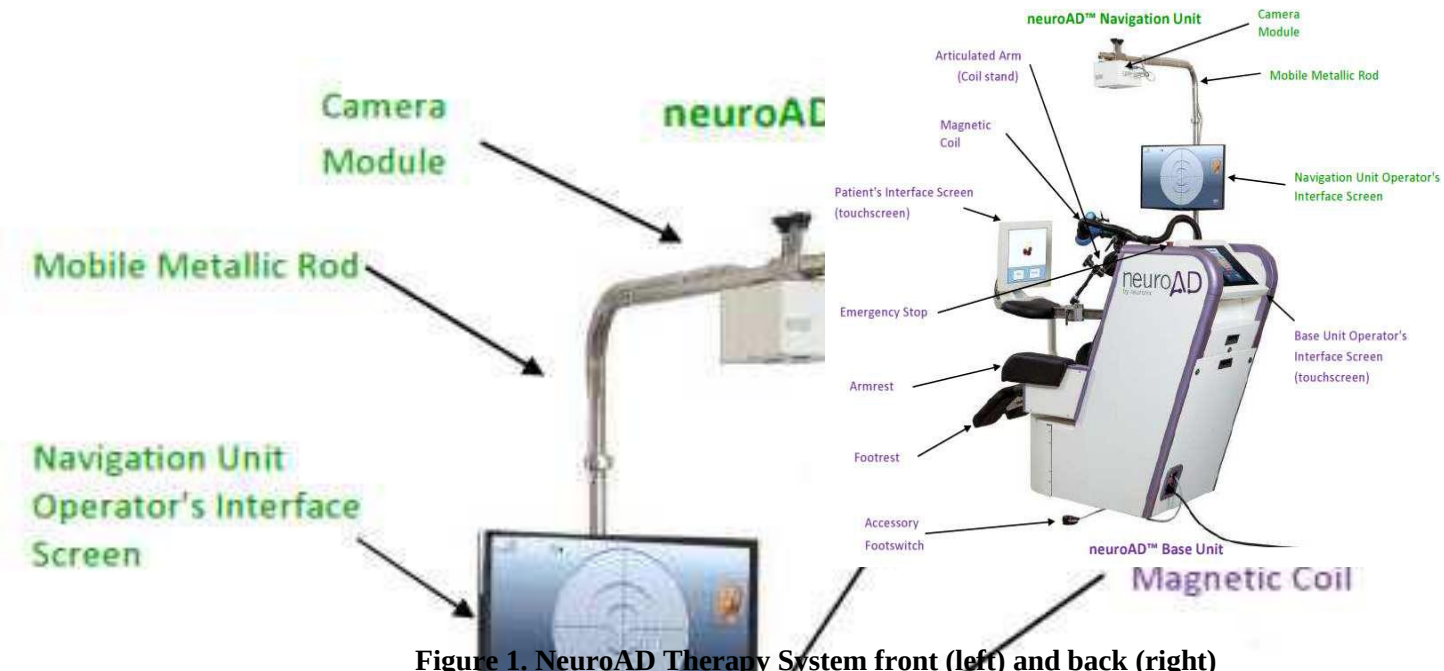
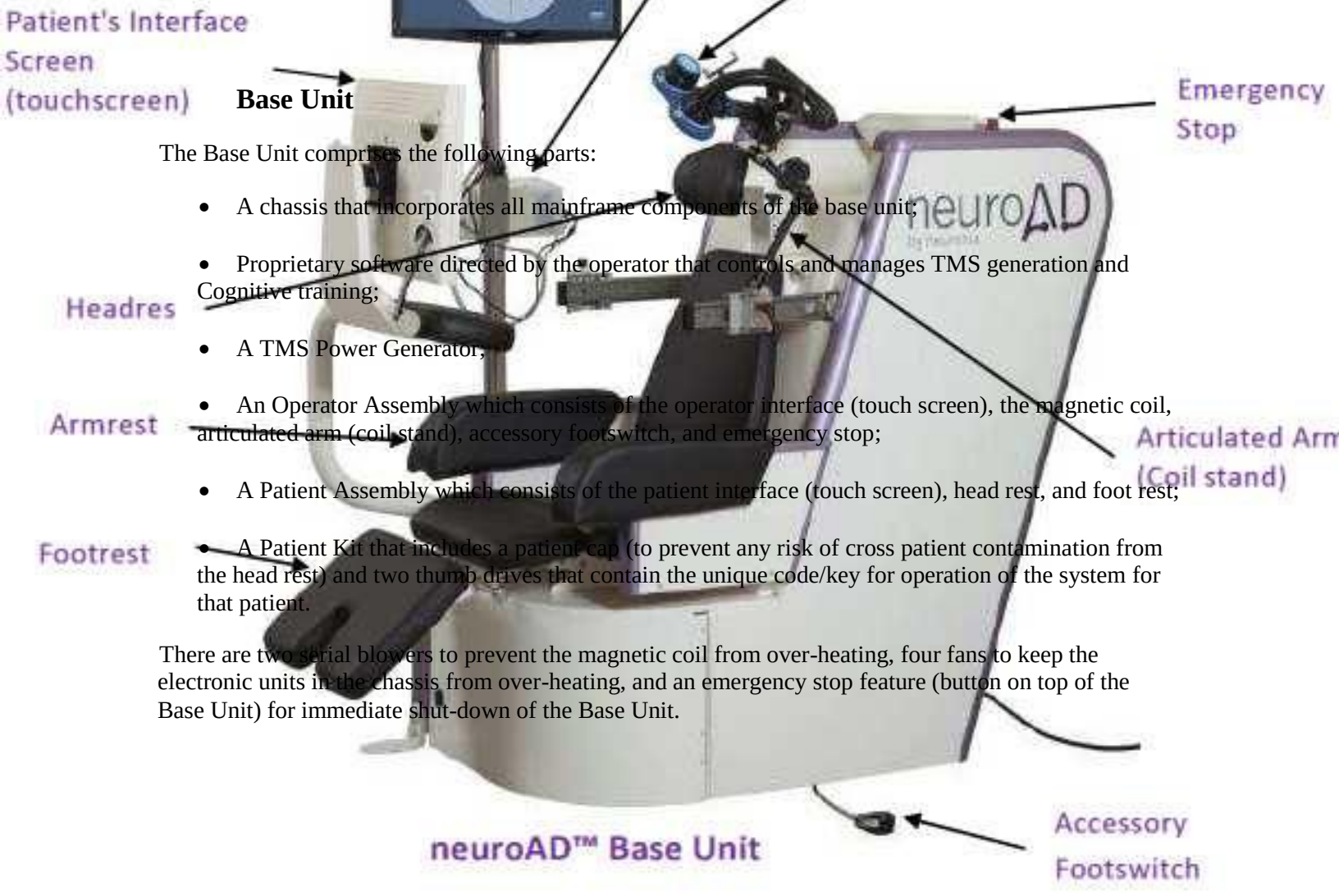


Figure 1. NeuroAD Therapy System front (left) and back (right)



Base Unit

The Base Unit comprises the following parts:

- A chassis that incorporates all mainframe components of the base unit;
- Proprietary software directed by the operator that controls and manages TMS generation and Cognitive training;
- A TMS Power Generator;
- An Operator Assembly which consists of the operator interface (touch screen), the magnetic coil, articulated arm (coil stand), accessory footswitch, and emergency stop;
- A Patient Assembly which consists of the patient interface (touch screen), head rest, and foot rest;
- A Patient Kit that includes a patient cap (to prevent any risk of cross patient contamination from the head rest) and two thumb drives that contain the unique code/key for operation of the system for that patient.

There are two serial blowers to prevent the magnetic coil from over-heating, four fans to keep the electronic units in the chassis from over-heating, and an emergency stop feature (button on top of the Base Unit) for immediate shut-down of the Base Unit.

Navigation Unit

The Navigation Unit comprises software, cameras, the PC, and accessories (e.g., pointer, head marker, and coil markers) that are used to ensure that the magnetic coil is positioned correctly on the patient's head over the target brain region. The treatment target areas of the brain are marked in advance based on the patient's MRI scan, and uploaded on the navigation unit using the patient's designated thumb drive.

The Optical Stereoscopic Module (i.e., the navigation functionality) is controlled by the proprietary software, which runs on the Navigation Unit PC.

Confidentiality of Patient Information

The thumb drives included in the Patient Kit are marked as "1-Lead" and "2-Slave" – the second drive is used for backup. All relevant patient data is kept on the patient thumb drives, which are kept confidential by the Operators. The Operators are responsible for backing up the records daily per the instructions in the User Manual. No patient information is stored locally on the Base Unit or Navigation Unit.

neuroAD Version History

The neuroAD used in the pivotal study and presented for consideration is the NICE-V3. However, data collected from an earlier version of the device, NICE-XP1, is also included in the de novo submission.

The NICE-XP1 was comprised of an earlier base unit design based on the Magstim Rapid-2 for TMS stimulation, connected to a controller and a user graphical display, (for computerized cognitive training and for receiving user feedback). This version used the Rogue Research navigation system and a couch seat. The NICE-XP1 was used in four of the six studies referred to as "Pilot Studies", e.g., in Table 3 of the original de novo. Three of the studies have published articles to reference (Bentwich et al., 2011; Rabey & Dobronevsky, 2016; Rabey et al., 2013) while the fourth is unpublished.

The review team has not identified any specific concerns that would affect our ability to leverage clinical data from the NICE-XP1. While the prior version of the system required manual switching between cognitive training levels (this process is now performed automatically by the software), the algorithm for modifying the difficulty of cognitive tasks remains the same. The two versions of the base unit provide identical TMS performance characteristics. Thus, the therapeutic effect of the two versions is expected to be the same.

Therapy Description

The following sections briefly describe the stimulation, training, and the protocol for treating a patient.

Transcranial Magnetic Stimulation (TMS) Therapy

The Magnetic Coil generates and applies a pulsed magnetic field to tissue immediately beneath the coil in response to commands from the processor to the power modules. The TMS stimulator consists of an electronic power supply chain, a single large capacitor for energy storage, a triggering component that allows the charge stored on the capacitor to discharge, and an inductive coil that receives the stored energy. TMS operates with the high current flowing through the coil, generating a magnetic field, which stimulates the targeted area of the brain. High-intensity current is rapidly turned on and off in the coil through the discharge of a capacitor, in the frequency of 10Hz. This produces a time-varying magnetic field that lasts for approximately 380 μs. The magnetic field strength is approximately up to 1.0 to 1.5

Tesla (approximately the same intensity as the static magnetic field used in clinical magnetic resonance imaging) at the surface of the coil – the strength of the magnetic field decreases with increasing distance from the coil.

The coil used in the neuroAD system (Magstim Rapid2 3530-00 70mm Double Rapid Air Coil) has been evaluated previously and is the same coil that was cleared for marketing via the 510(k) process for the stimulation of peripheral nerves (see K051864) and also for the treatment of depression (see K143531). It is not manufactured by Neuronix. It is an enclosed copper coil in a figure-8 shape that is mounted on the coil gantry and placed against the patient's head.

The technical characteristics and operating parameter range of the device are shown in the tables below:

Table 3. TMS System Characteristics

Power Specifications	
Voltage	100-240VAC
Current max	10A Momentary 20A/2sec
Frequency	50/60 Hz
System Outputs	
Maximum Operating Current (Ip2p)	8000 A
Maximum Operating Voltage (Vp2p)	2800V
Maximum Operating Frequency	10Hz
Coil Characteristics	
Coil Name	Magstim 3530-00 70mm Double Rapid Air Coil (K051864)
Configuration	Biphasic Figure 8 Coil
Number of Windings	2
Inside Diameter of each Winding	46.75 mm
Outside Diameter of each Winding	86.15 mm
Number of Turns per Winding	11
Connecting Cable Length	1.75 m
Pulse waveform	Biphasic Single Cycle
Pulse period	380 µs
Weight	3.9 kg
Frequency	10 Hz
Peak Voltage Reached	1200 V @ 100% power
Maximum Acoustic noise	80 dBA
Core Material	Air

The following table provides specific information regarding the intended therapy parameters for the neuroAD which are a subset of the available system parameters that are listed above. These are described further in the U.S. Treatment Protocol section below.

Table 4. neuroAD TMS therapy parameters

Output Stimulation Parameters:	
(b)(4)Proprietary Information	
Magnetic Field:	
(b)(4)Proprietary Information	

Cognitive Training

The Cognitive Training tasks are presented on a computer medical touch screen, mounted as part of the neuroAD system, located about 50cm from the patient. The users select their answers by touching graphical buttons on the touch screen according to the specific task.

The cognitive training is conducted with the proprietary software, per a pre-set algorithm. The software reads/writes previous treatment data from/to dedicated thumb-drives, which are provided for each patient. The thumb-drives are initialized for each patient on the first treatment session. The software runs the Cognitive Training paradigms. Dedicated tasks were developed to activate six cortical brain regions known to be affected by AD, to be performed by patients in parallel to TMS administration. The tasks are summarized in Table 5 below.

Table 5. Brain Regions and Cognitive Training Paradigm

Targeted Brain Region and Description	Corresponding neuroAD Cognitive Training Tasks
Broca Syntax and grammar tasks	• Sentence Similarity • Right/Wrong Sentences
Wernicke Comprehension of lexical meaning and categorization tasks	• Words/Pseudo Words • Categories
Dorsolateral Prefrontal Cortex Left & Right (DLPFC L&R) Action naming, object naming and spatial memory of shapes and colors	• Color (no Location) • Location (no Color) • Location and Color • Action Naming • Subject Naming • Word Recall
Somatosensory Cortex Left and Right (Parietal L&R) Spatial attention for shapes and letter	• Rectangles-Red • Rectangles-Blue • Letters B • Letters T • Letters M

For example, when the Wernicke area is trained, the patients need to name the group the object belongs to by touching the screen, shown in Figure 2 below.

Figure 2. Cognitive Training Example - Wernicke

The cognitive training protocol is a set algorithm that progresses by both task complexity and time. The Cognitive Training tasks were developed with a scale of difficulty levels permitting each patient to advance through the levels at an individually appropriate pace. The difficulty levels were developed by controlling for task variables (e.g., number of objects, time required to complete the task, etc.). All programs have multiple difficulty levels. The level automatically updates based on performance according to an algorithm developed by the sponsor.

The cognitive tasks presented by the NeuroAD system for each brain region correspond approximately with the neuroanatomical structures either proven or purported to assess these cognitive domains. These tasks have equivalent paper-and-pencil neuropsychological tests and assess the major domains known to be impaired in Alzheimer's disease.

U.S. Treatment Protocol

Pre-treatment, the spatially discrete regions of the cerebral cortex (i.e. treatment area) and facial anatomical landmarks are identified and marked on a standard structural MRI head scan. The four anatomical landmarks serve as registration points during the procedure (tip of nose, left ear, right ear, and bridge of nose). These marks are noted as spatial coordinates on a computerized file which is uploaded to the patient's thumb-drive.

For the Calibration (coil) and Registration (patient landmarks) processes, the camera of the Navigation Unit uses infrared LED illumination to identify and calibrate with pre-specified markers. When the patient is situated on the Base Unit, the Operator uploads the patient's marked MRI and applies the Patient's Head Marker (headband with tripod sensors) to the patient's head. During the registration process the operator points the pointer instrument at the patient's anatomical landmarks to co-register with the digital MRI landmarks. The operator follows a coil calibration process that is similar to the registration process by using the pointer instrument directed at the center of the coil.

By calibrating the magnetic coil and registering the patient's facial anatomical landmarks, the neuroAD™ Navigation Unit identifies the magnetic coil location relative to the specific treatment area. The neuroAD™ Navigation Unit tracks the magnetic coil's location (using the magnetic coil marker) relative to the specific treatment area (using the Patient Head Marker) throughout the procedure and displays the distance on the screen. Thus, the operator can adjust and reposition the magnetic coil as required. The operator manually navigates the magnetic coil to the specific brain region by maneuvering the Articulated Arm of the Navigation Unit. The magnetic coil is in the intended location when indicated by a big green

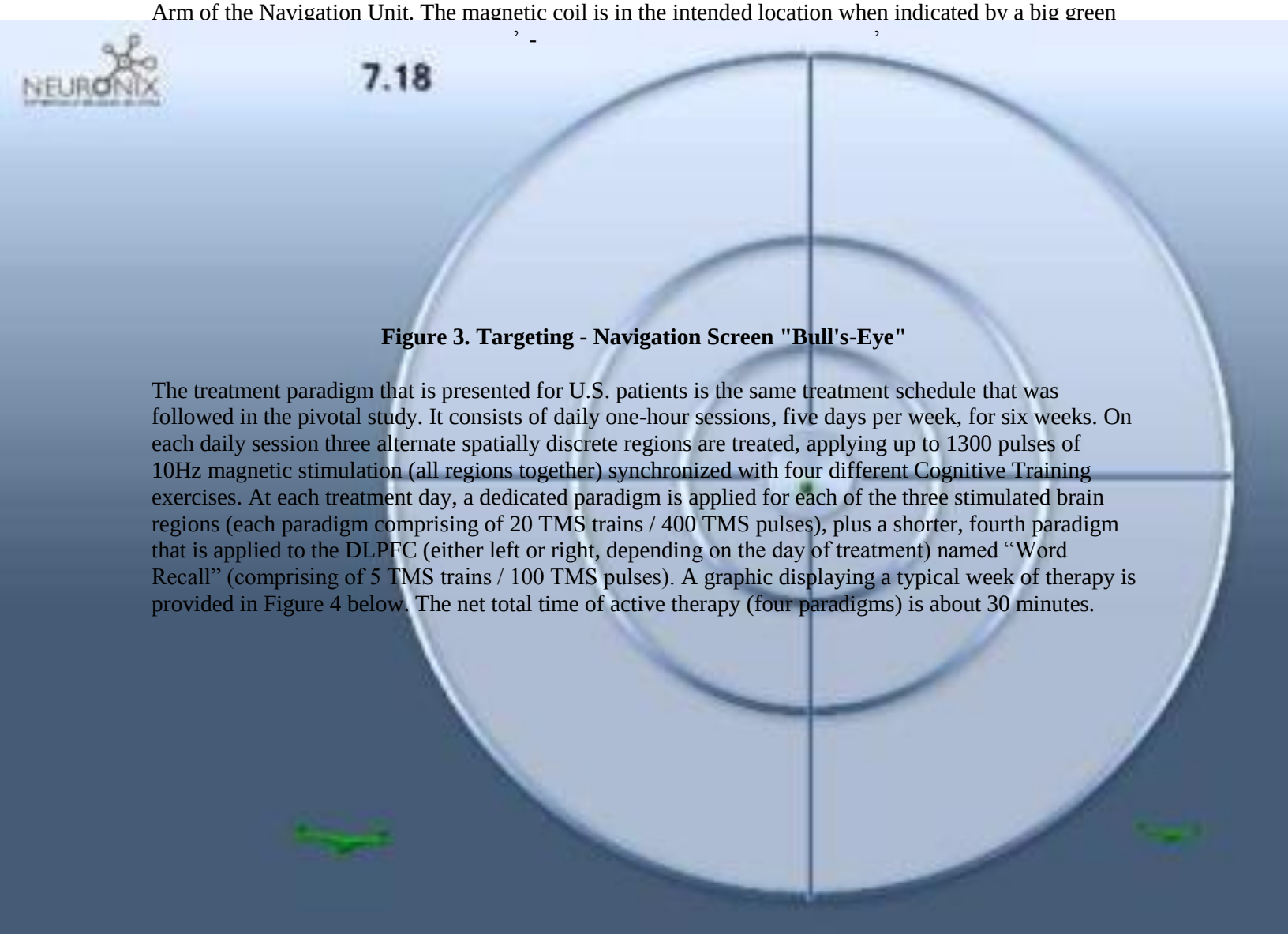


Figure 3. Targeting - Navigation Screen "Bull's-Eye"

The treatment paradigm that is presented for U.S. patients is the same treatment schedule that was followed in the pivotal study. It consists of daily one-hour sessions, five days per week, for six weeks. On each daily session three alternate spatially discrete regions are treated, applying up to 1300 pulses of 10Hz magnetic stimulation (all regions together) synchronized with four different Cognitive Training exercises. At each treatment day, a dedicated paradigm is applied for each of the three stimulated brain regions (each paradigm comprising of 20 TMS trains / 400 TMS pulses), plus a shorter, fourth paradigm that is applied to the DLPFC (either left or right, depending on the day of treatment) named "Word Recall" (comprising of 5 TMS trains / 100 TMS pulses). A graphic displaying a typical week of therapy is provided in Figure 4 below. The net total time of active therapy (four paradigms) is about 30 minutes.

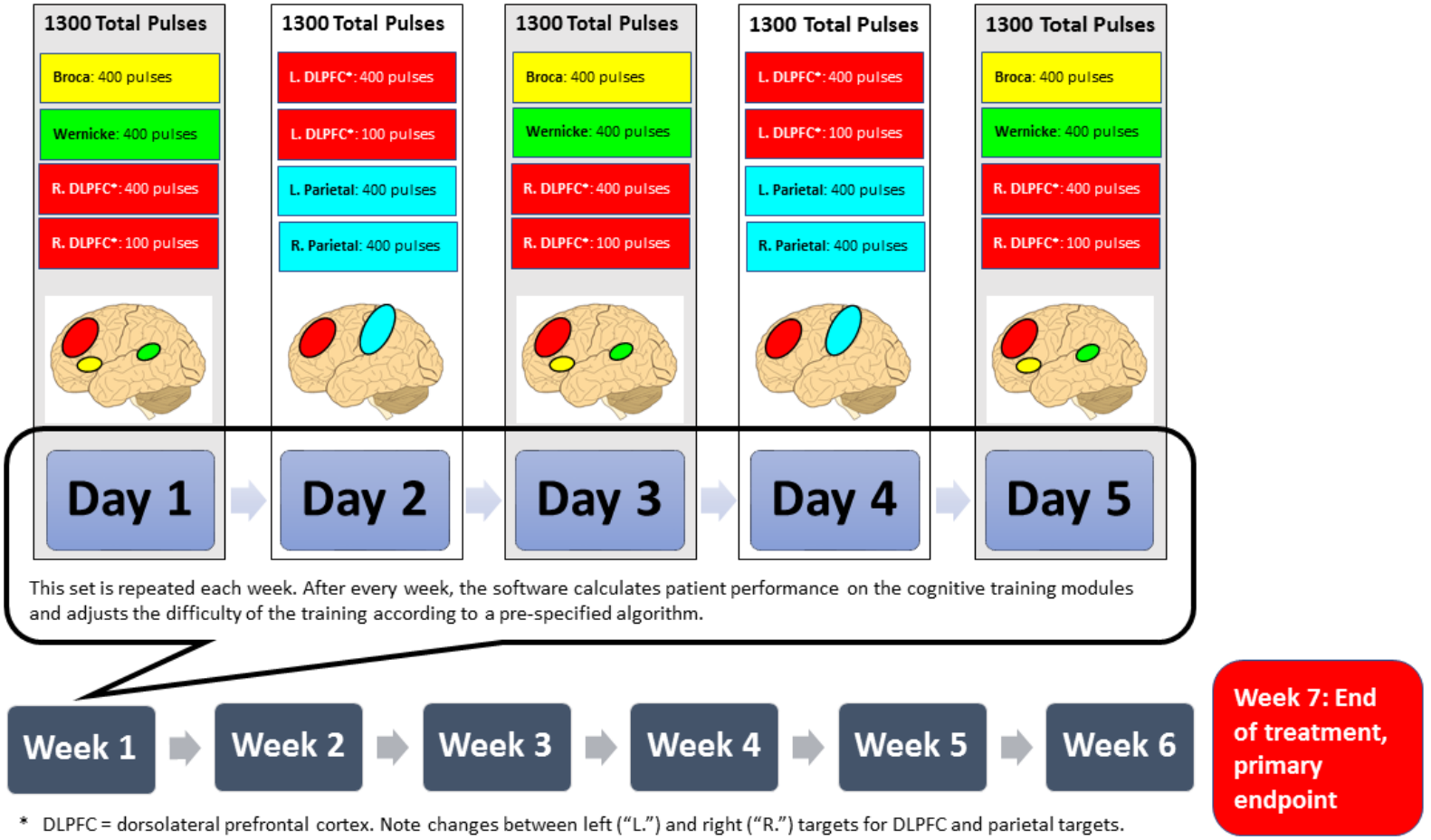


Figure 4. neuroAD Treatment Paradigm

As discussed in section above, patients perform the associated cognitive tasks in conjunction with the cortical stimulation by the TMS. The patient practices demo Cognitive Training tasks prior to commencing of the actual Cognitive Training paradigm. The purpose of the demo is to assure the patient understands the instructions and nature of the task.

A typical treatment session is as follows:

1. Patient's data is uploaded to the neuroAD Base Unit and Navigation Unit via the thumb drive.

2. Cognitive Training Module accesses previous treatment information, determines treatment day number, and sets the brain regions and cognitive paradigms.

3. Navigation Unit accesses patient's MRI file, magnetic coil is calibrated, patient's anatomical landmarks are registered to the system.

4. The stimulation intensity is set using the standard procedure described in Schutter et al., 2006, i.e., single TMS pulses are applied while Operator monitors the patient for motor reaction. The intensity will be set at 90-110% of the patient's recorded motor threshold (MT) on the day of treatment; Broca - 90% of MT, Wernicke and Dorso Lateral left and right- 100% of MT, and Parietal left and right-110% of MT.

5. Treatment is started – (note that the Operator can stop the treatment at any time by pushing the Emergency Stop which shuts down the neuroAD completely by stopping the electrical supply)

    a. Operator navigates magnetic coil to specific region using the neuroAD Navigation Unit (based on software guidance). When in place, the Operator locks the coil in place and starts the treatment stimulation and cognitive training paradigm; the progression of the therapy is then automatically managed by the Base Unit software.

    b. Each Cognitive Training paradigm starts with presenting the patient with simple instructions on the upcoming Cognitive Training exercise. The operator confirms the patient understands the instructions, and if needed, provides a few example questions using the neuroAD User Interface. Next, the Operator starts the treatment and continuously monitors the magnetic coil location on the neuroAD Navigation Unit Operator Interface Screen.

    c. The paradigm stops automatically upon completion and the software user interface guides the Operator on the next treatment area/paradigm

    d. When four paradigms are completed the treatment sessions ends.

6. The device software calculates patient performance and adjusts the difficulty level of the Cognitive Training for the following treatment sessions.

Note regarding the treatment protocol of the pivotal study: The impact of increasing or decreasing the difficulty level of the cognitive training in the active group throughout the study was not a variable that was prespecified for assessment. Therefore, this information has not been requested by or presented to FDA. The impact of increasing or decreasing the cognitive training on the results of the pivotal study is not clear. Please refer to the Pivotal Clinical Study Design section for more information about the design of the pivotal study.

Regulatory History

Neuronix, Ltd had 4 presubmission interactions with FDA before submitting the de novo application. The FDA submission number, approximate date of feedback, and a brief summary of topics covered is provided in Table 6 below.

Table 6. Neuronix, Ltd Presubmission Interactions

FDA Submission	Date of Feedback	Description
(b) (4)		

The sponsor submitted the subject de novo, DEN160053, to FDA on November 14, 2016 and the submission was reviewed. This initial review concluded with a letter requesting Additional Information on March 13, 2017. While the request included nonclinical and clinical topics, the clinical request is the focus of this meeting. Details of the clinical Additional Information Request are provided below.

The clinical dataset intended to support the marketing application was the 130-patient pivotal study. At the primary endpoint the sham group showed a greater negative change on the ADAS-Cog from baseline to 7-weeks (a negative change suggests improvement on the assessment scale) than the neuroAD treatment group with a difference between groups favoring sham of +1.45 points (p=0.09). At the 12-week secondary ADAS-Cog assessment the trend reversed from 7 weeks and the result favored the active group by -0.42. There was no neuroAD therapy or sham therapy provided between 7 weeks and 12 weeks.

A test for interaction between baseline ADAS-Cog and treatment group was pre-specified in the primary analysis at 7 weeks, and a significant interaction was found between baseline ADAS-Cog and treatment group at 7 weeks. A post-hoc analysis was performed on the baseline ADAS-Cog $\leq$30 subgroup at both 7 and 12 weeks. This removed 8 subjects from the active group and 8 subjects from the sham group at baseline. The largest change in the ADAS-Cog from baseline in favor of the neuroAD treatment group over sham was -1.61 points and was found in the ADAS-Cog$\leq$30 subgroup at the 12-week endpoint. At

the 7 week endpoint, in the post-hoc subgroup of subjects with ADAs-Cog≤30 the result favored the sham group (+0.47), as in the overall cohort.

Regarding the CGI-C secondary endpoints, in the overall population, CGI-C mean scores indicated a difference between groups of 0.02 in favor of treatment at 7 weeks and 0.35 in favor of treatment at 12 weeks. In the ADAS-Cog≤30 subgroup, the CGI-C at the 7-week time point resulted in a difference of 0.07 in favor of the treatment group. At the 12-week timepoint, the 0.40 difference between groups was the highest magnitude CGI-C difference in favor of the treatment group.

Given these results, and the proposal to limit the indication to a post-hoc subgroup, ODE did not believe the evidence submitted demonstrated a clinically or statistically meaningful benefit and a new clinical study to confirm the post-hoc results of the pivotal study ([redacted]).

The sponsor submitted a request to meet and discuss the request for additional clinical information (a "Submission Issue" meeting, [redacted]), which was received on April 19, 2017. The submission outlined Neuronix's plans to attempt to address the deficiency by providing additional published clinical information in lieu of designing and conducting a new clinical study.

Neuronix responded to ODE's March 13, 2017, request for additional information on October 6, 2017. The supplement included additional clinical data from outside the U.S. (including data from a Korean pivotal study, which included only mild Alzheimer patients with baseline ADAS-Cog≤30) and additional analyses of both the data previously submitted to FDA as well as the new clinical data sources.. The sponsor also included a Physician Survey to support the clinical meaningfulness of the clinical data results and letters of support (n=5) from physicians. More information about the Supplemental Investigations and the Physician Survey is found in the Appendices of this document.

The review team consulted with FDA's Network of Experts in March-April 2018. The Network of Experts is a vetted network of partner organizations and their member scientists, clinicians and engineers who can provide CDRH and CDER staff with rapid access to expertise when it is needed to supplement existing knowledge and expertise within the Centers. In this case, ODE sought to obtain additional external input on general (i.e., non-confidential) information regarding the ADAS-Cog scale and clinical meaningfulness of changes on this scale.

After review of the additional information, ODE continued to have concerns regarding the effectiveness of the device which led to a denial decision delivered to Neuronix on June 22, 2018. Neuronix appealed the denial decision in a letter dated July 23, 2018 and requested supervisory review. The supervisory review concluded that CDRH would benefit from additional external scientific and clinical perspective on whether the data in the submission demonstrate that the probable benefits of the device outweigh the probable risks. The file was re-opened and referred to the Neurological Devices Advisory Panel to further discuss the evidence submitted to the agency before CDRH renders a final decision

Pivotal Clinical Study

For ease in interpretation, the data in charts and figures in the sections below will generally adhere to the following color conventions:


Study populations combining active and sham subjects


Study subjects receiving active treatment


Study subjects receiving sham treatment


Difference between active and sham groups for the mean ADAS-Cog change from baseline for a particular visit (7 or 12 weeks). Green indicates a difference favoring the active group, red indicates a difference favoring the sham group.

In cases where the figure includes multiple data points for the same group (e.g., data from the active group at two points in time), the difference between the data will be distinguished by a pattern.

Pivotal Clinical Study Design

The following section contains information on the pivotal study that was presented for the neuroAD Therapy System. Of the studies described in this summary and contained in the submission, the pivotal study is the largest, most rigorously designed, and contains elements that reduce uncertainty in the results (e.g., a prespecified statistical analysis plan which defined the analysis populations and endpoints, and including a blinding assessment, a plan for handling missing data, poolability testing for study sites both in the United States (US) and outside the United States (OUS), a sensitivity analysis, and a covariate analysis)

The pivotal study was a prospective, randomized, multi-center, double-blind, sham-controlled study. It was designed for up to 150 subjects and 10 study sites (in the US and Israel). The sample size was expected to provide greater than 90% power for demonstrating primary efficacy.

The pivotal study was designed with FDA input through presubmissions (b) (4) , and Q140479. The Agency initially recommended that Neuronix study the effects of the cognitive training and the TMS components separately. Neuronix chose not to pursue this course, and FDA notified Neuronix that there were no major concerns regarding the final proposed clinical study design.

The neuroAD Therapy System treatment may be administered in conjunction with pharmaceutical treatment and was studied as such in the pivotal trial. For the purpose of assessing the neuroAD Therapy System performance only, subjects were required to be on stable dose of AD drugs at least 60 days prior to joining the study and throughout the study as detailed in the inclusion criteria.

Pivotal Study Analysis Groups

The pivotal study was designed to include two groups randomized in a 2:1 Treatment Group to Sham Group ratio. While there are two distinct "active" therapy components, TMS and Cognitive Training, the study did not assess these components separately. The two groups are described below:

Treatment: Active TMS+Active Cognitive Training
Sham Control: Sham TMS+Sham Cognitive Training

Table 7. Pivotal Study Active and Control Groups

	Active	Sham
TMS	As described in section above.	Sham TMS is delivered by the same device as the real TMS. The coil will be placed at three regions per day, similar to the real treatment, and the machine creates identical noise as real TMS. The only difference is that no electromagnetic energy is produced or delivered (sound recording only).
Cognitive Training	As described in section above.	Subjects will receive "pseudo cognitive treatment" on the same interface as is used by the active group. Patients will choose on a touch screen "like" or "don't like" when presented with pictures. In addition, the patients in the Sham group will watch short nature (or other) movies, as well as being shown picture slides, without being asked specific questions.

Eligibility Criteria

Patients could be included in the study only if they met all the following inclusion criteria.

1. Male or female age 60-90 years
2. Patients diagnosed with mild or moderate stage of Alzheimer's Disease, according to the DSM-IV criteria
3. MMSE score 18 to 26
4. ADAS-Cog above 17
5. Physical clearance for study participation as evaluated by the clinician
6. Spouse, family member or professional caregiver agree and capable of taking care for the participation of the patient in the study (answering questions regarding the patient's condition and assuming responsibility for medication)
7. Agreement to participate in approximately 15 weeks during the study
8. Normal to near-normal vision and hearing with correction as needed (e.g., corrective lenses, hearing aid)
9. Fluent in English or Hebrew
10. Minimum of 8th grade education
11. If medicated for AD, then use of cholinesterase inhibitors, Memantine or Ginko-biloba for at least 3 months and on stable dose for at least 60 days prior to screening and during the course of study (including follow-up period)

Patients were excluded from the study for any of the following reasons:

1. Clinical Dementia Rating (CDR) of 0, 0.5, or 3
2. Severe agitation
3. Mental retardation

4. Patient lacking capacity to consent to study participation (this condition may be removed in accordance with local State regulations and IRB approval)
5. Unstable medical condition
6. Use of benzodiazepines or barbiturates 2 weeks prior to screening
7. Pharmacological immunosuppression
8. Participation in a clinical trial with any investigational agent within 6 months prior to study enrollment
9. History of Epileptic Seizures or Epilepsy
10. Contraindication for performing MRI scanning
11. Contraindication for receiving TMS treatment according to a TMS questionnaire
12. Pregnant women and women who have the ability to become pregnant unless they are on an acceptable method of contraception during the study
13. Patients with depression, bipolar disorder or psychotic disorders or any other neurological or psychiatric condition (whether now or in the past), which the Investigator finds as interfering with the study
14. Alcoholism or drug addiction as defined by DSM-IV within last 5 years (addicted more than one year and or in remission less than 3 years) or severe sleep deprivation
15. Patients with metal implants in the head, (i.e. cochlear implants, implanted brain stimulators and neurostimulators, aneurysm clips) with the exception of metal implants in mouth
16. Patients with personal history of either any clinically defined medical disorder (which the Investigator finds as interfering with the study) or any clinically defined neurological/psychiatric disorder (other than AD), including (but not limited to): stroke, brain lesions, substance abuse, vitamin B12 deficiency, abnormal thyroid function, cerebrovascular condition, other neurodegenerative disease, head trauma, multiple sclerosis; or personal history of previous neurosurgery or head trauma that resulted in loss of consciousness (unless the investigator confirms the disorder to be irrelevant to the study)
17. Patients with any signs or symptoms of increased intracranial pressure, as determined in a neurological exam
18. Cardiac pacemakers
19. Implanted medication pumps
20. Intracardiac lines
21. Significant heart disease
22. Currently taking medication that lower the seizure threshold
23. Patients on which TMS Motor Threshold cannot be found
24. Patients who underwent TMS treatment in the past

Sample Size Calculation

Sample size considerations were based on demonstrating superiority of the neuroAD to Control on the change in ADAS-Cog from baseline to week 7. Based on data collected in the early US and Israeli trials, Neuronix estimated mean Change to be -4.6 and 0 in Active and Sham respectively, with standard deviation (SD) of about 4.0 in each group. The effect size based on these previous studies was estimated by (-4.6-0)/4.0= -1.15.

A sample size of 17 subjects per group was intended to provide 90% power to demonstrate a difference between the two groups. Power is based on an independent groups t-test with two-sided Alpha = 0.05. Thus, a total of 34 subjects was needed in this trial for requisite power. Adding 10% to account for dropout, Neuronix obtained a total of 38 subjects needed for both the Active treatment and Sham groups.

During pre-IDE interactions, we (FDA) recommended that the study include 100 randomized subjects. Adding 2 roll-in patients at each site, the total number of subjects in this trial was calculated to be up to

120 subjects (50 per arm, plus up to 2 roll-in patients receiving active treatment), to achieve over 90% power to demonstrate superiority of Active Treatment to Control on change from baseline to end of treatment on ADAS-Cog.

As noted in Appendix I. Nonclinical Studies due to a software error, a concern was raised that some of the subjects (specifically, two Active subjects) received cognitive training at non-optimal difficulty level; hence, it was decided that these subjects will be excluded from the efficacy analyses blind to their outcome. Consequently, up to 30 additional patients were added to the trial. Additionally, the randomization rate change was changed effective January 10th, 2015 to a ratio of Treatment to Sham of 2:1, to have at least 50 valid randomized subjects in each study arm as recommended by FDA. Thus, the sample size was increased to be up to 150 subjects to participate in the trial.

Primary Effectiveness and Safety Populations

Subjects were randomized in a 2:1 treatment group to sham group ratio. In addition, the first two recruited patients at each site were part of a "run-in / roll-in" phase (not randomized) and assigned to the active treatment group.

Primary Safety Population
The safety population included all patients for whom active or sham treatment was initiated, including roll-in patients (if Visit Date at the week 1 day 1 visit is not missing).

Primary Effectiveness Population (PE)
The primary efficacy population included all randomized subjects who had at least one baseline measurement on the primary efficacy endpoint ADAS-Cog and who participated in at least one post-baseline treatment visit (Active or Sham) (i.e., if visit date for at least one of the post baseline visits is available).

Included were subjects with no major entry violations as determined by blinded review, listing of major entry violations were provided by the sponsor.

Roll-in subjects were not included in the primary efficacy population (as defined as subjects marked as "not applicable" on the randomization form).

Per Protocol Population (PP)
The per-protocol (PP) population will be a subset of the primary efficacy population of subjects who had no major protocol violations likely to affect outcome, and who:

- Had at least 24 treatment visits of the planned treatments (i.e. at least 24 of visit dates for the planned treatments are not missing); and
- Did not miss more than 2 visits in any week of the planned six weeks of treatment (i.e. at least 3 visit dates are not missing for each week of the planned treatment); and
- Did not miss two visits (two or more) during more than 2 weeks of treatment, of the planned six weeks treatments.

Listing of major protocol deviations likely to affect outcome was provided by the sponsor.

Study Schedule and Follow-up

Figure 5 below provides a representation of the study schedule and assessments.

It is important to note that no active or sham treatment was delivered after Week 6. There was no "maintenance therapy" proposed or delivered during the pivotal trial (Appendix II. Supplemental Investigations Study Designs section of this memo). Endpoints assessed at 7 weeks were approximately 1-week post-treatment and endpoints assessed at 12 weeks were approximately 6 weeks post-treatment.

Additionally, please note that the neuroAD Therapy System treatment may be administered in conjunction with pharmaceutical treatment. For the purpose of assessing the neuroAD Therapy System performance only, subjects were required to be on stable dose of AD drugs at least 60 days prior to joining the study and throughout the study as detailed in the inclusion criteria. Concomitant medications were monitored throughout the study at each study visit.

Item	Screening	Baseline		Active/Sham Treatment Phase						Follow-Up	
Visit Number	1	2	R	3-7	8-12	13-17	18-22	23-27	28-32	33	34
Scheduled Week	-21 to day -1	-14 to day -4		Wk 1	Wk 2	Wk 3	Wk 4	Wk 5	Wk 6	Wk 7	Wk 12
Informed consent	X										
Medical History & Demographics	X										
Physical Exam	X										
Concomitant Medications	X			X	X	X	X	X	X	X	X
MMSE	X										
ADAS-Cog (Blinded Rater)		X								X	X
ADCS-CGI-C (Blinded Rater)		X								X	X
CDR	X										
TMS-Safety Questionnaire	X										
TMS-Motor Threshold Measurement		X									
MRI		X (3 workings days turnaround)									
Randomization			X								
TMS-Motor Threshold Measurement – daily (Active group only)				X	X	X	X	X	X		
neuroAD Treatment (Daily) (Group 1&2)				X	X	X	X	X	X		
Adverse Events		X		X	X	X	X	X	X	X	X
Inclusion & Exclusion Criteria	X	X (Review)									
Study Deviation	X	X		X	X	X	X	X	X	X	X
Study Exit Group 1											X
Study Exit Group 2											X
Treatment Group 1 – neuroAD System Active (real) Treatment. Patient #1-#2 assigned to the treatment group.											
Treatment Group 2 – neuroAD System Sham Treatment.											

Figure 5. Pivotal Study Schedule of Events

Screening Window (Visit 1)
The screening window was up to 21 days before first treatment administration. During the screening visit all patients provided either written informed consent or written essence of consent prior to participation in any study related procedures. For patients who had a legal authorized representative (LAR) at the time of consent, which is frequently necessary for patients with Alzheimer's Disease, the LAR provided written informed consent for the patient's participation in the study prior to any study related procedures. Per local requirement, the investigator assessed the patient's capacity to provide informed consent, and if incapacity was found, the LAR caregiver was required to provide written consent. Upon protocol amendments and/or informed consent updates, an updated informed consent was obtained from patients, LARs and caregivers, as required.

The following were collected or performed before any study-specific procedures were performed:

- Informed Consent Form
- Complete physical examination, neuropsychological assessment, current medications that subject is using, and medical and surgical history. Ensure that all inclusion/exclusion criteria are met
- Cognitive Dementia Rating (CDR)
- Mini Mental State Examination (MMSE)
- Safety TMS screening questionnaire (Keel et al. 2001)
- Participants were instructed not to be engaged in any other experiment and/or new therapy treatment during their participation in the study

Only patients who met all eligibility criteria were enrolled and continued to the Baseline visit.

Baseline Window (Visit 2)
The baseline window was up to 14 days before first treatment administration. The following activities were performed during the baseline visit:

- Clinical Global Impression of Change (CGI-C)
- Alzheimer's Disease Assessment Scale - Cognitive (ADAS-Cog)

> **Note:** Please see Clinical Context section above for more information about the ADAS-Cog and the CGI-C

The CGI-C rater and ADAS-Cog rater were two different, independent raters, each of whom was blinded to the other's ratings. Patients who were not scored within the inclusion criteria of ADAS-Cog (> 17) were not referred to the Motor Threshold measurement and MRI procedure and were considered screening failures. Reason for exclusion was recorded in the study screening log.

- TMS-Motor Threshold Measurement: An initial measurement of TMS motor threshold was performed to determine the MT and to confirm the patient's eligibility to receive TMS intervention. Standard procedure was followed (Schutter et al. 2006). If the subject's Motor Threshold was not identified, he/she was removed from the study as a screen failure.

- MRI-Scan: All subjects underwent a structural MRI-scan to identify excluded disorders including non-Alzheimer brain pathology (analysis provided by investigator) and mark brain areas to be treated. No contrast media agent was used for the MRI scan. Patients who did not meet all of the eligibility criteria including the MT threshold and brain scan criteria were not randomized and were excluded from the study as screen failures.

Subjects were considered enrolled once they completed all screening and baseline activities including Baseline MRI scan. The subjects were randomized after the screening and baseline activities were complete (i.e., the baseline ADAS-Cog and CGI-C assessments were conducted *before* the subjects were randomized).

Details regarding the Active/Sham Treatment Phase (Visits 3-32) are provided above.

Pivotal Study Endpoints

Primary Safety Endpoint

- Adverse events (AE) including serious AE (SAE) occurring at any time during the trial or follow up, whether or not deemed related to study device.

Primary Effectiveness Endpoint

- Change in ADAS-Cog from Baseline to week 7 (visit 33)

Secondary Endpoints

The secondary endpoints are as follows:

- Change in ADCS-CGIC from Baseline to week 7 (visit 33).
- Change in ADCS-CGIC from Baseline to week l2 (visit 34 - 6 weeks after discontinuation of treatment)
- Change in ADAS-Cog from Baseline to week 12 (visit 34 - 6 weeks after discontinuation of treatment)

Study Blinding Procedures and Blinding Assessment

The following involved parties were blinded to the subjects' group assignment until the end of the study: Investigators, ADAS-Cog & ADCS-CGI-C raters, subjects, caregivers, sponsor's team, and the clinical research organization ("CRO") management team.

The following parties were unblinded to the subjects' group assignment: neuroAD Operators (members of the sites' teams), the CRO monitors, and the CRO IVRS team. No other parties were unblinded during the course of the study.

To assure proper blinding throughout the study, only the subject and operator were present in the treatment room during the procedure; source worksheets were clearly marked and kept out of reach of unauthorized personnel. In addition, the eCRF was password protected and audit trail was maintained. Blinding regarding treatment information and subjects' group assignment was maintained throughout the study until database lock.

Blinding Assessment

A blinding assessment questionnaire was conducted for the patient/caregiver at the end of the first week of treatment (Visit 7) and the clinical raters after the completion of the six weeks of treatment (Visit 32).

Patient/Caregiver (Visit 7): At the end of the first week of the Active/Sham treatment phase (end of fifth active/sham treatment session) the patient and caregiver will be presented separately with the following:

- Patient: "Do you know if the treatment that you received today was an actual treatment or a placebo treatment?"
- Caregiver: "Do you know if the treatment the patient received today was an actual treatment or a placebo treatment?"

Possible responses are: "Actual Treatment", "Placebo" or "Not sure / cannot tell".

The results of the blinding assessment questionnaire that was conducted for the patient/caregiver are shown using percentage of subjects/caregivers reporting each possible response.

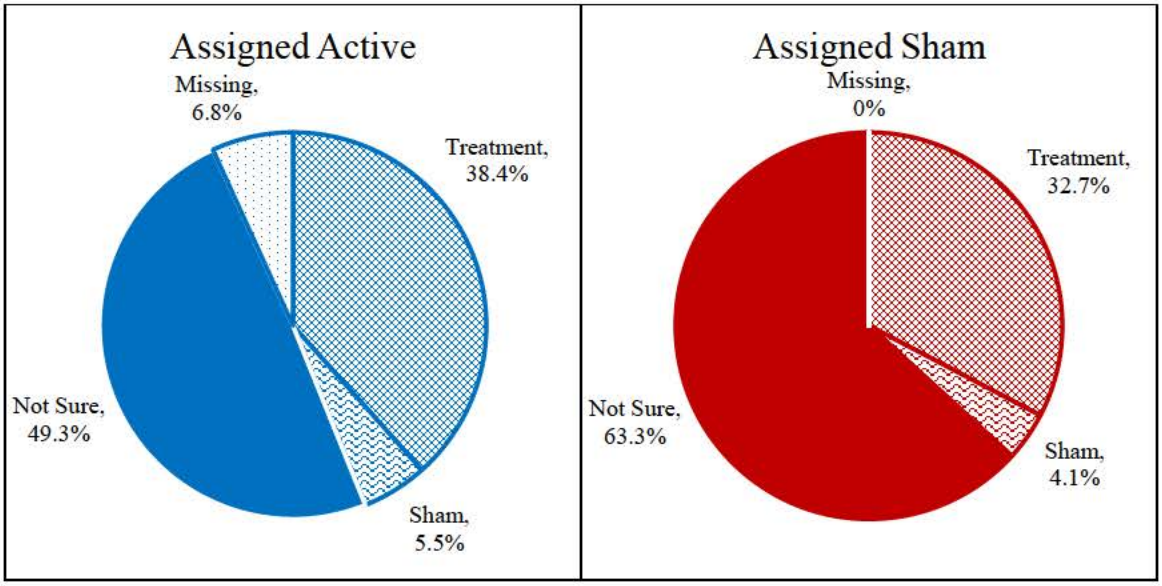


Figure 6. Pivotal Study - Blinding Assessment, Subject Responses (%)

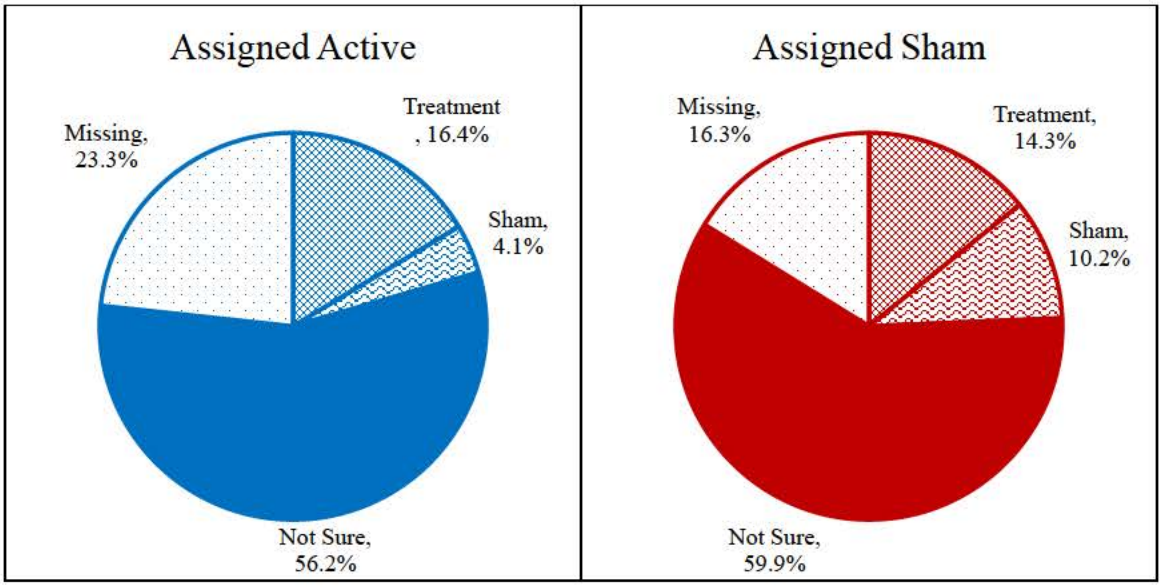


Figure 7. Pivotal Study - Blinding Assessment, Caregiver Responses (%)

ADAS-Cog and CGI-C Raters (Visit 32): After completion of six weeks of active/sham treatment, and prior to first follow-up (visit 33) ADAS-Cog and CGI-C raters were presented separately with the following question

- "Do you know if the treatment the patient received was an actual treatment or a placebo treatment?"

Possible responses were: "Actual Treatment", "Placebo" or "Not sure / cannot tell".

The results of the blinding assessment questionnaire that was conducted for the clinical raters are described below.

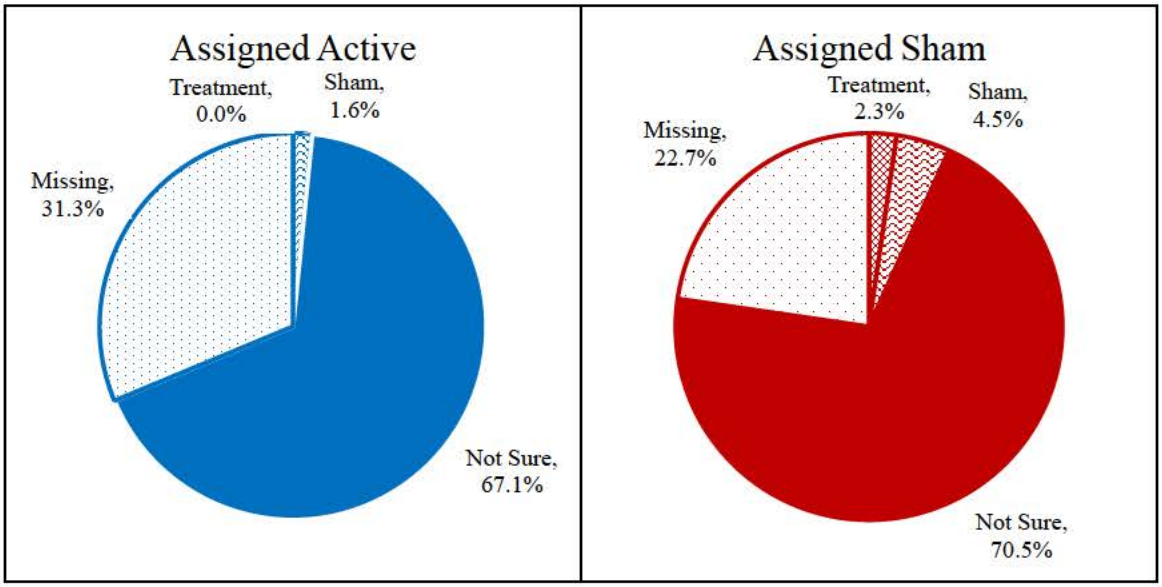


Figure 8. Pivotal Study - Blinding Assessment, ADAS-Cog Rater Responses (%)

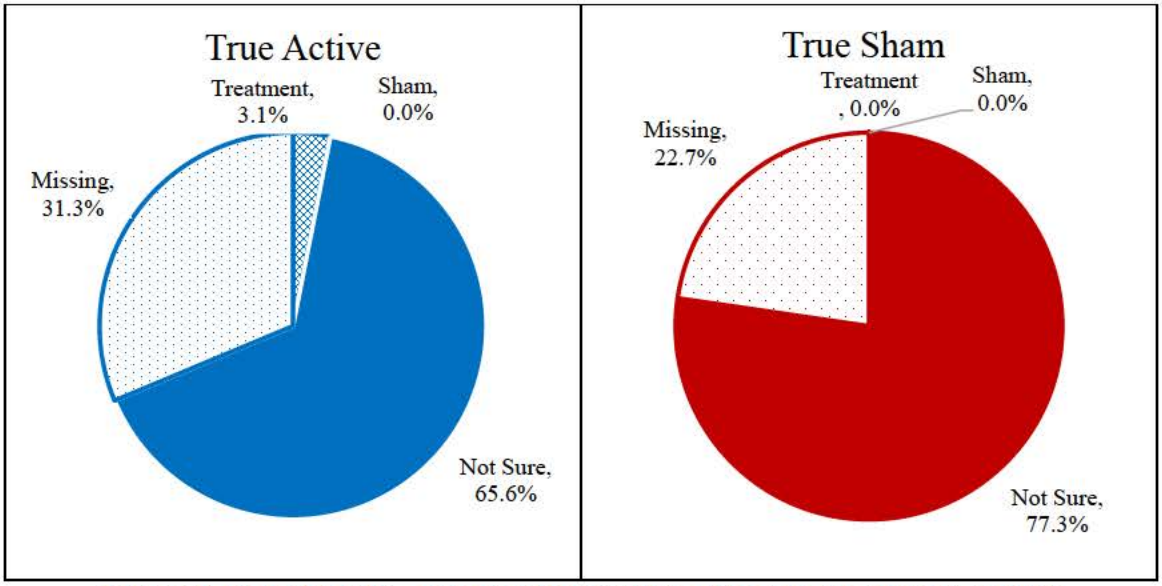


Figure 9. Pivotal Study - Blinding Assessment, CGI-C Raters Responses (%)

Sensitivity and Covariate Analysis

A sensitivity analyses was included in the statistical analyses plan for the primary efficacy endpoint of change in ADAS-Cog from Baseline to Week 7. It was specified that analyses assessing the possible impact of missing primary data should include best case, worst case, and tipping point. As specified in the SAP, missing data for the primary and secondary endpoints were imputed using multiple imputation methodology. This imputation was carried out and yielded results very similar to those obtained based on the non-imputed data, hence the results described in this report are based on observed (non-imputed) data only. Additional analyses were carried out based on observed data only as well

As outlined in the study protocol (Section 9.6), ADAS-Cog at baseline was specified as a covariate to be assessed in the primary analysis. The possible effect of various baseline covariates was tested by assessing the interaction term in the following model: Change = $\text{ADAS-Cog}_{\text{Baseline}}$ + Covariate + Group + Covariate x Group

The following covariates were tested:

- Global CDR
- Age
- Gender
- Ethnicity
- Race
- BMl
- Nicotine use
- Center
- Disease Severity (mild vs. moderate, based on MMSE)*
- Level of Education
- Hand dominance
- Medicated for AD
- Time since AD diagnosis to treatment
- Number of missed treatment visits
- Baseline TMS-Motor Threshold value (%) – this covariate was not pre-specified but was explored post-hoc**

In the clinical study report, the sponsor did not provide results of the interaction tests but stated that covariate analyses showed no significant impact on the ADAS-Cog results recorded at Week 7 and Week 12, with the exception of the *baseline* ADAS-Cog assessment. In the section FDA Summary Comments on Pivotal Study Post-Hoc Analysis ADAS-Cog≤30 Subgroup, FDA presents its own analyses of interactions between treatment and baseline ADAS-Cog.

***Note for pivotal study: Disease Severity**
MMSE was used to stratify randomization in the sponsor's study by disease severity. Baseline MMSE did not interact with treatment group at 7 weeks ($p = 0.35$) or at 12 weeks ($p = 0.89$).

****Note for pivotal study post-hoc analysis, ADAS-Cog≤30 subgroup: Baseline Motor Threshold**
In the original submission the sponsor hypothesizes that a finding in a post-hoc covariate analysis regarding the interaction between treatment group and Baseline MT may lend support for indicating the device to the subgroup of patients with a baseline ADAS-Cog score ≤30. The sponsor states that better effectiveness outcomes are associated with higher MT, which is associated with a lower baseline ADAS-Cog score. In the de novo submission, the sponsor presented Figure 10 below which has been copied

here. The figure shows ADAS-Cog change from baseline by baseline MT, with different colors for treatment group. Least squares lines were fit to each treatment group scatterplot, but it appears that the lines are influenced by a few outlying observations. Otherwise, there appears to be no apparent relationship between MT and treatment effectiveness.

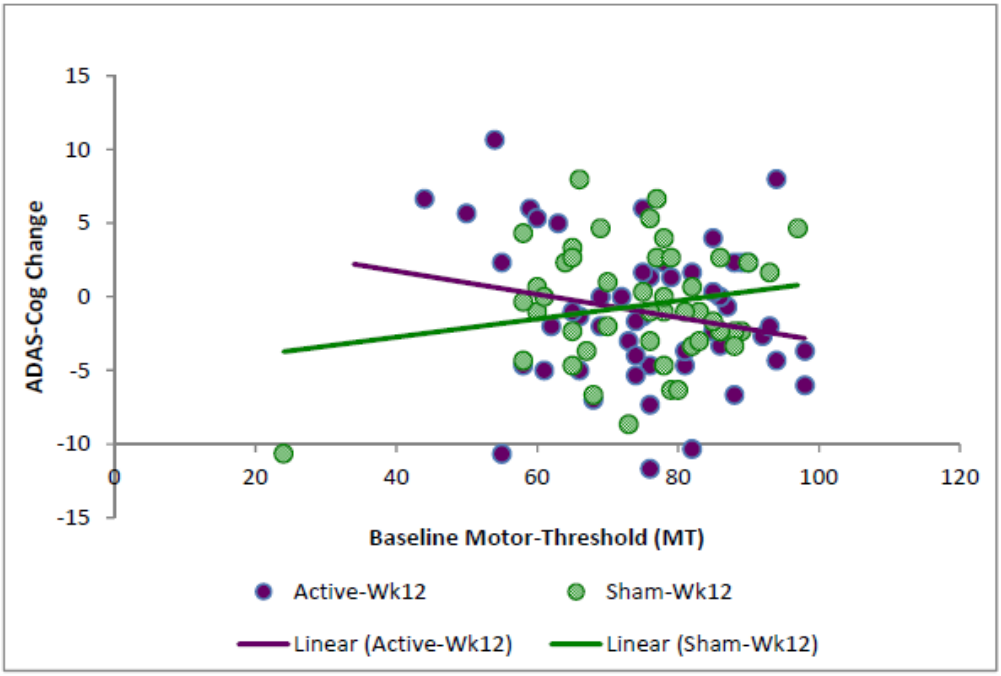


Figure 10: ADAS-Cog Change on Week 12 by Baseline TMS-Motor Threshold (Primary Efficacy Population)

Figure 11 was also presented by the sponsor in the original submission to demonstrate a correlation between baseline ADAS-Cog and MT. The correlation is -0.405. FDA superimposed a horizontal line at baseline ADAS-Cog = 30. The MTs above that line (with a range from about mid-40% to high 70% – ignoring the outlier at about 24%) are somewhat lower than the MTs below the line (mid-50% to 100%). But, from Figure 10 above, it is not clear that higher MT results in better response to neuroAD over sham. Therefore, a correlation between baseline MT and baseline ADAS-Cog does not appear to explain why subjects with lower baseline ADAS-Cog responded better to neuroAD over sham, but that those with higher baseline ADAS-Cog did not.

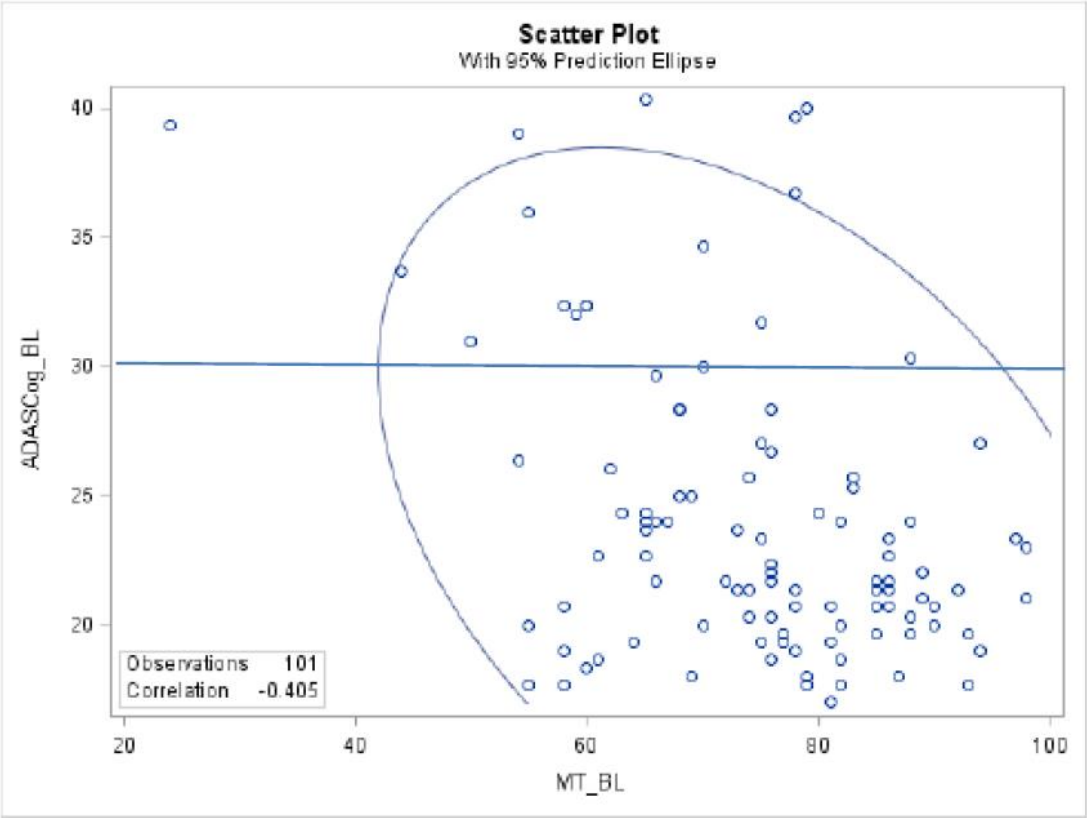


Figure 11: Correlation between baseline ADAS-Cog and motor threshold

Study Site Training Activities

A study initiation visit was performed at each of the sites prior to any study related procedures.

Study initiation visit included:

- Protocol training – presenting the technology & study rational, reviewing study objectives, endpoints, arms, procedures, visits, etc.
- Hands-on device training – including both Active and Sham procedures

Initiation was performed by a Sponsor representative, often with the presence and participation of the Clinical Research Organization (CRO) representative, and lasted between 4-5 days per site. All study team members at the site were required to participate in the visit, as per their role, including investigators, coordinators, raters and device operators. In the absence of one of the site team members, another training session was scheduled.

Operators were required to complete hands-on training that lasted several days as stated above. Training included overview of the technology and device, followed by a detailed demonstration of the procedures and hands-on practice. Sponsor representative was also present at the site during first procedures (run-in subjects) to support and guide the operators in case additional questions or issues were raised.

Study raters were trained on the different scales used in the study (ADAS-Cog, CGI-C & MMSE) using a dedicated training program developed by the sponsor and Dr. (b) (6) the (b) (4)

(b) (4) [redacted]. Dr. (b) (6) [redacted] also served as the instructor for the training sessions performed conveniently via WebEx.

(b) (4) [redacted], (the CRO appointed on monitoring the study) was responsible for training the coordinators and investigators on EDC data entry, using the Medidata dedicated training program. Quality checks for the EDC were performed by (b) (4) [redacted] team. All study cases were monitored and verified against source data. All informed consent forms were monitored as well. Study monitors were trained on study specific items by the Sponsor.

At the end of the study, all ADAS-Cog forms (scale used to evaluate study primary endpoint) were monitored in a blinded manner by the sponsor, without knowing subject's randomization, to assure accuracy and consistency. Also, a short discussion/interview was conducted with each of the CGI-C raters (scale used to evaluate study secondary endpoint) to assure similar administration of the scale was applied.

All study sites were audited by the sponsor (Regulatory Binders and Informed Consents), to assure compliance with regulatory requirements. Subjects' records were not reviewed during the audit visits, to maintain sponsor representatives blinding.

Pivotal Clinical Study Results

As the ADAS-Cog results are reported as a ***change*** from baseline, figures displaying changes in ADAS-Cog do not include all possible values (the scale has a maximum of 70 points). Figures displaying changes in ADAS-Cog are generally shown with a range of -5 to +1 for consistency; however, several individual figures use a different range due to the magnitude of the change exceeding -5 points.

Figures displaying CGI-C scores are all shown on a number line from 1-7, which are the possible values.

Subject Enrollment and Disposition

The pivotal study enrolled the 130 subjects across 10 sites. Each of the ten clinical centers had a local Principal Investigator (PI) with overall responsibility for the study at the site, two raters who performed the assessment scales (ADAS-Cog, CGI-C), a study coordinator, and at least two Operators who operated the neuroAD Therapy System during procedures (lead Operator and back-up Operator). At some of the sites the study coordinator also served as an Operator of the neuroAD Therapy System. It is important to note that the two raters who performed the assessment scales (ADAS-Cog & CGI-C) were independent from study staff that met the subjects throughout the study. The raters were blinded to the subject's group assignment and also to each other. Site information is displayed in Table 8. Subject enrollment per site is shown graphically in Figure 12 below.

Table 8. Pivotal Study Site Information

Site #	Site Name	Principal Investigator
101	Lou-Ruvo Center for Brain Health, Cleveland Clinic, Las Vegas, NV	Charles Bernick, MD, MPH
102	Banner Sun Health Research Institute, Sun City, AZ	Marwan Sabbagh, MD
103	NYU Langone Medical Center, New York, NY	Steven H. Ferris, PhD Stella Karantzoulis, PhD
104	Palm Beach Neurology and Premiere Research Institute, West Palm Beach, FL	Carl Sadowsky, MD
105	Cleveland Clinic, Cleveland, Ohio	Babak Tousi, MD
106	Beth Israel Deaconess Medical Center, Harvard, Boston, MA	Alvaro Pascual-Leone, MD, PhD
107	Miami Jewish Health Systems, Miami, FL	Marc Agronin, MD
108	ATP Clinical Research, Costa Mesa, CA	Gustavo Alva, MD
109	Roskamp Institute, Sarasota, FL	Andrew P. Keegan, MD
201	Asaf-Harofe Hospital, Beer-Yakov, Israel	Carmel Armon, MD

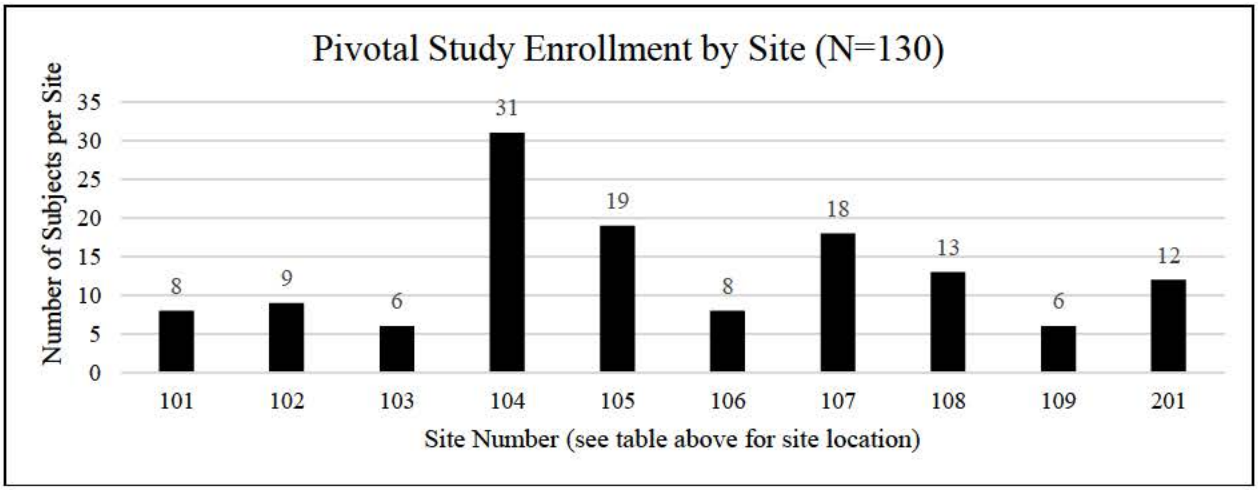


Figure 12. Pivotal Study Enrollment by Site

One hundred and thirty-one (131) subjects completed screening and baseline procedures and were found eligible to participate in the study. However, one subject experienced an unrelated SAE that required hospitalization prior to randomization and treatment initiation (upper respiratory infection and cervical fracture due to a fall), and therefore was withdrawn from the study before actually entering the study. Therefore, 130 subjects were enrolled in the study. Of the 130 enrolled subjects, 20 subjects were considered as run-in/roll-in subjects and are only included in the safety analysis; leaving 110 subjects that were randomized to receive either active or sham treatment.

FDA developed the flowchart below to summarize the subject disposition information that was provided in the clinical study report of DEN160053.

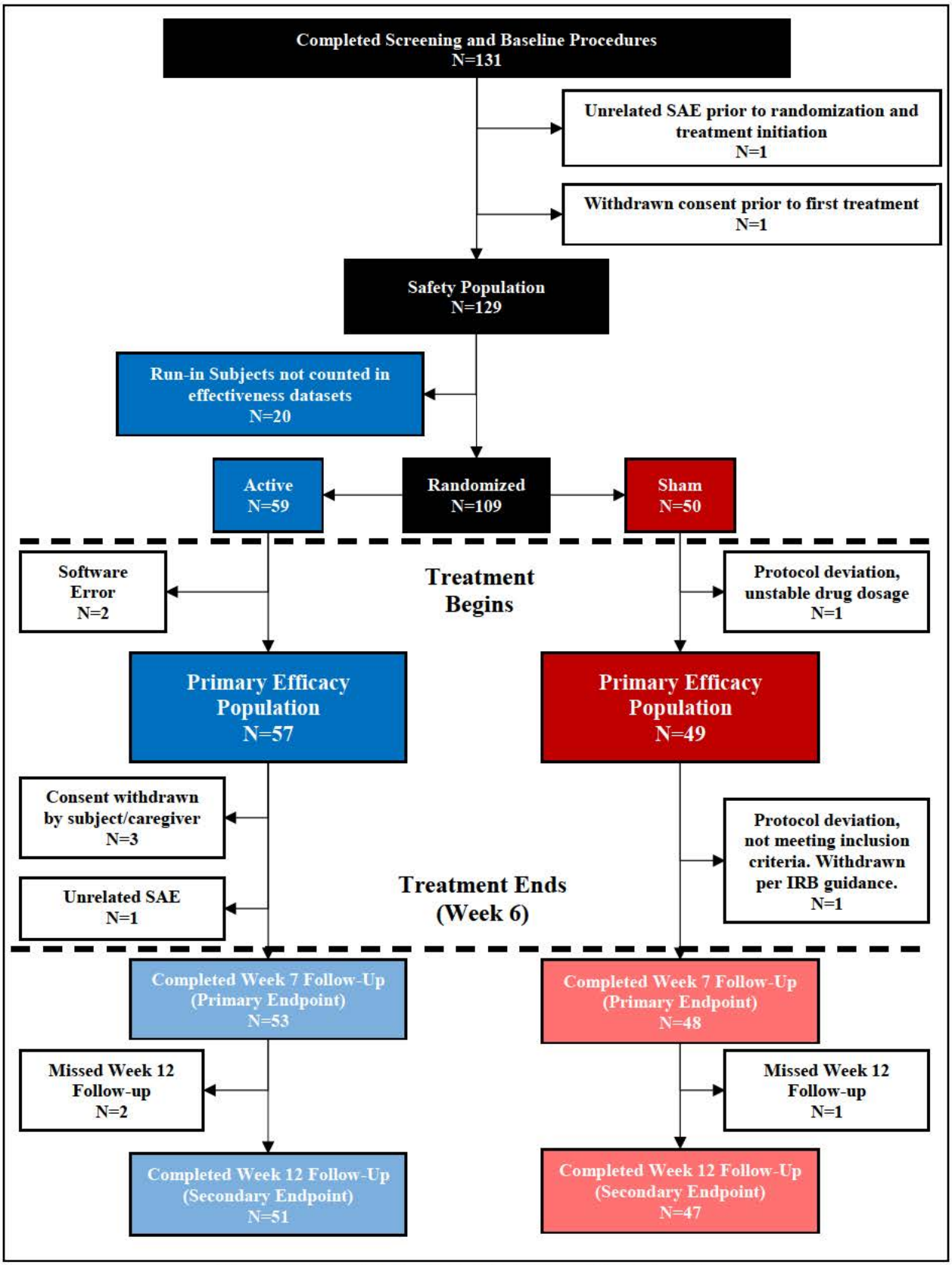


Figure 13. Pivotal Study Subject Disposition Flowchart

Safety Population: 129 subjects out of the 130 enrolled (99.2%) were included in the safety analysis population, including the 20 run-in subjects. One subject was excluded from the safety analysis as the subject withdrew consent post randomization and prior to first treatment session administration (subject did not receive full or partial active or sham treatment).

Primary Efficacy Analysis Population: The Primary Efficacy Population (PE) is comprised of randomized subjects who had at least one baseline measurement on the primary efficacy endpoint ADAS-Cog and who participated in at least one post-baseline treatment visit (Active or Sham).

Included were subjects with no major entry violations as determined by blinded review. Also, as indicated above, subjects that may have been affected due to the software error were excluded from the Primary Efficacy Population and blind to their outcome (but were included in the Safety Population).

Roll-in subjects (n=20) were not included in the Primary Efficacy Population.

Overall 106 out of the 129 subjects included in the Final Safety Population were included in the Primary Efficacy Population (96.3% out of the randomized subjects); this included 57 active subjects and 49 sham subjects. Twenty-four subjects were excluded from the Primary Efficacy Population for the following reasons:

- Twenty roll-in subjects (two first subjects enrolled at each site)
- Two subjects that may have been affected by the software error (both active group)
- One subject for whom treatment was not initiated (withdrawn consent before randomization)
- One subject who had major entry violation – increase in AD drug dosage (an increasing dose of AD drug may bias the results in favor of the investigational device) (sham group)

Of the 106 subjects in the PE population, 101 subjects completed the week 7 follow-up (53 active and 48 sham) and 98 subjects completed the week 12 follow-up (51 active and 47 sham).

In the active group, six subjects were lost from baseline to week 12. From baseline to week 7, three subjects/caregivers withdrew consent and 1 subject experienced an unrelated SAE. From baseline to week 12, two subjects missed the week 12 visit.

In the sham group, 2 two subjects were lost from baseline to week 12. From baseline to week 7, one subject was withdrawn per IRB guidance for a protocol deviation and one subject missed the week 12 follow-up visit.

Per-Protocol Population: The Per-Protocol Population (PP) is a subset of the Primary Efficacy Population which consisted of subjects who had no major protocol violations likely to affect outcome, and who (as defined in the protocol):

- Had at least 24 treatment visits of the planned treatments; and
- Did not miss more than 2 visits in any week of the planned six weeks of treatment; and
- Did not miss two visits (two or more) during more than 2 weeks of treatment, of the planned six weeks treatments.

Overall 98 subjects (50 active, 48 sham) were included in the Per-Protocol Population and 95 subjects (48 active and 47 sham) completed through the week 12 visit. All subjects that were excluded from the Per-Protocol Population were excluded for not meeting the minimum required visits as set forth above. The reasons for not meeting the minimum number of visits is detailed below:

- One subject death (reported SAE, see description below)
- Three subjects in the treatment group withdrew consent. Two subjects withdrew consent after one treatment session and one subject withdrew consent after 7 treatment sessions for an undisclosed reason.
- One subjects was withdrawn per sponsor request for not meeting the baseline ADAS-Cog eligibility criteria
- Three subjects missed three treatment sessions on the same week

Note: The PE population was the pre-specified analysis population for the primary effectiveness endpoint. Per the study protocol, the primary analysis was to be repeated using the Per Protocol (PP) population to ensure consistency when protocol violations were eliminated from the PP data set. All secondary analyses were to be conducted on both the PE and PP populations.

A summary table with rationale for early termination by analysis group is provided in Table 9 below.

Table 9. Reasons for Premature Termination of Study Subjects in Pivotal Study

Analysis Population / Study Group		Reasons for Premature Termination						All
		Subject and/or Subject's Legally Acceptable Representative Withdrew from the Study	Refusal of the Subject's Study Partner / Caregiver to Continue Follow- up Observation	Serious Adverse Event	Significant Protocol Deviation	Decision Made by the Investigator	Other	
		N	N	N	N	N	N	N
Safety	Treatment Group (N=79)	3	0	1	0	0	1	5
	Sham Group (N=50)	0	1	0	1	0	0	2
	All (N=129)	**3**	**1**	**1**	**1**	**0**	**1**	**7**
Primary Efficacy	Treatment Group (N=57)	3	0	1	0	0	1	5
	Sham Group (N=49)	0	1	0	1	0	0	2
	All (N=106)	**3**	**1**	**1**	**1**	**0**	**1**	**7**
Per Protocol	Treatment Group (N=50)	0	0	0	0	0	1	1
	Sham Group (N=48)	0	1	0	0	0	0	1
	All (N=98)	**0**	**1**	**0**	**0**	**0**	**1**	**2**

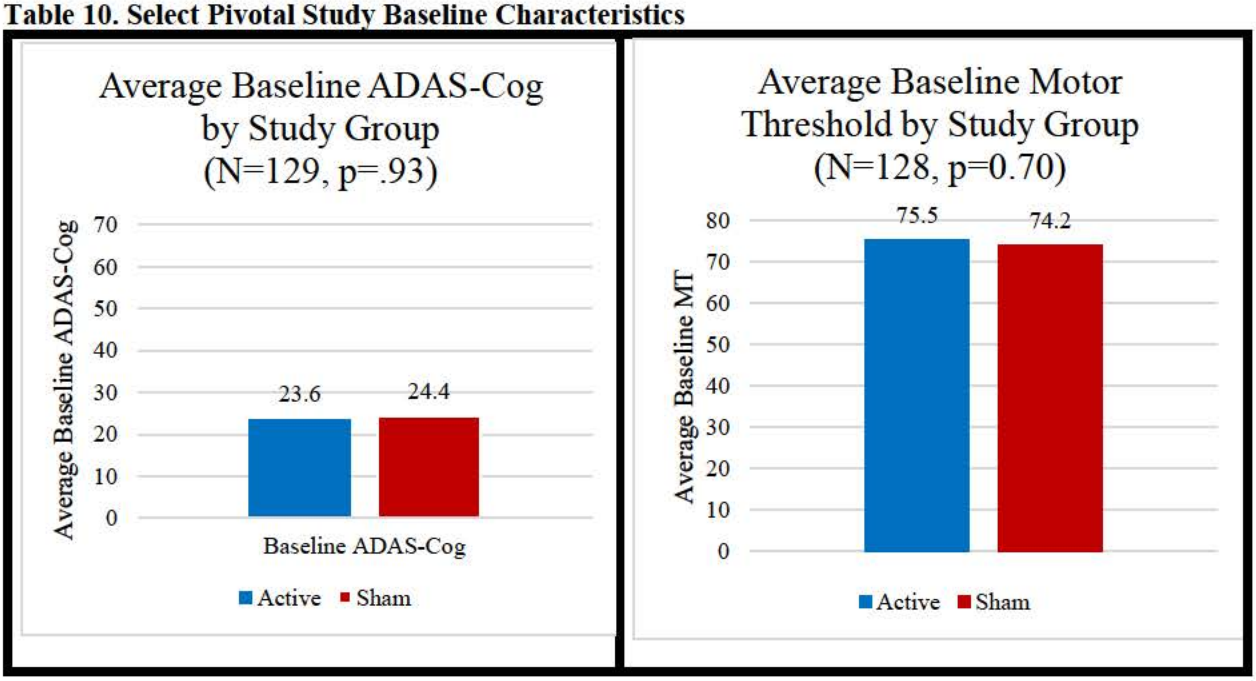
Baseline and Procedural Characteristics

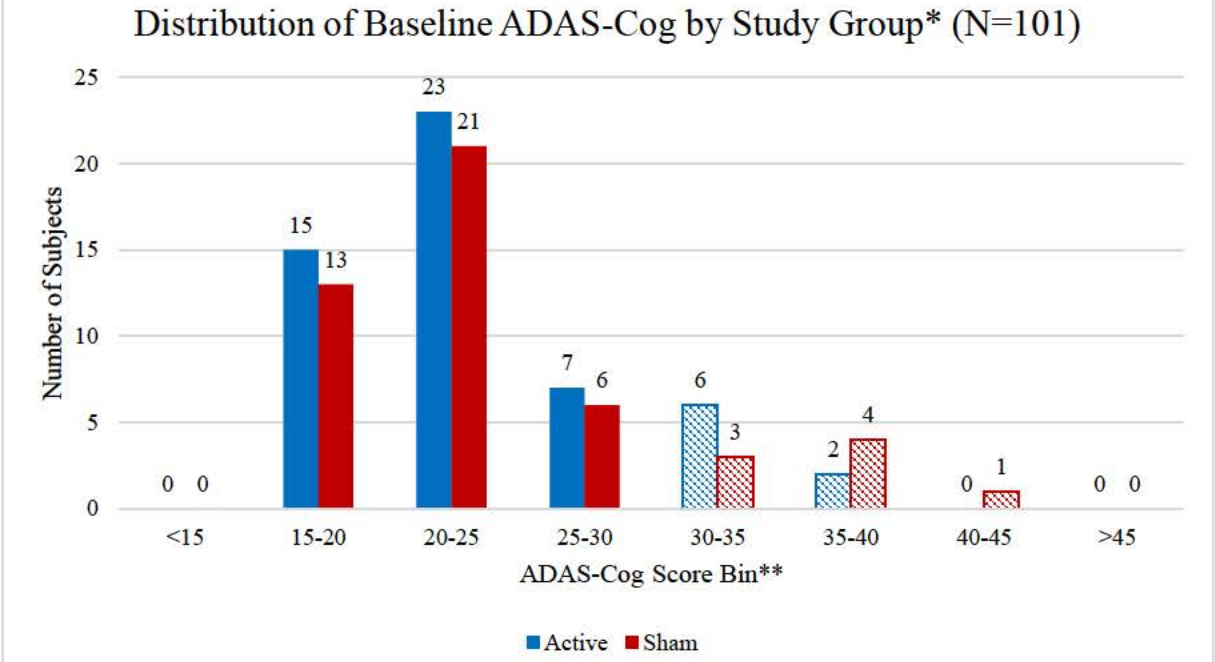
The following tables and charts describe the demographic and baseline characteristics of the subjects in the pivotal study.

Please note that the first two subjects at every site (20 subjects total) were treated as "run-in" subjects. These subjects were unblinded and are within the demographic data as "active subjects". The run-in subjects were not included in the efficacy analysis populations but were included in the safety population. This was pre-specified for the pivotal study.

For baseline information the FDA believes is important to illustrate the baseline characteristics between active and have we have provided the data by study group. However, please keep these 20 unblinded subjects in the active group in mind when looking at the demographic data broken down into active and placebo groups. Because we do not have concerns regarding the balance between group on any of the baseline characteristics, for other demographics we show the entire population, not broken into active or sham

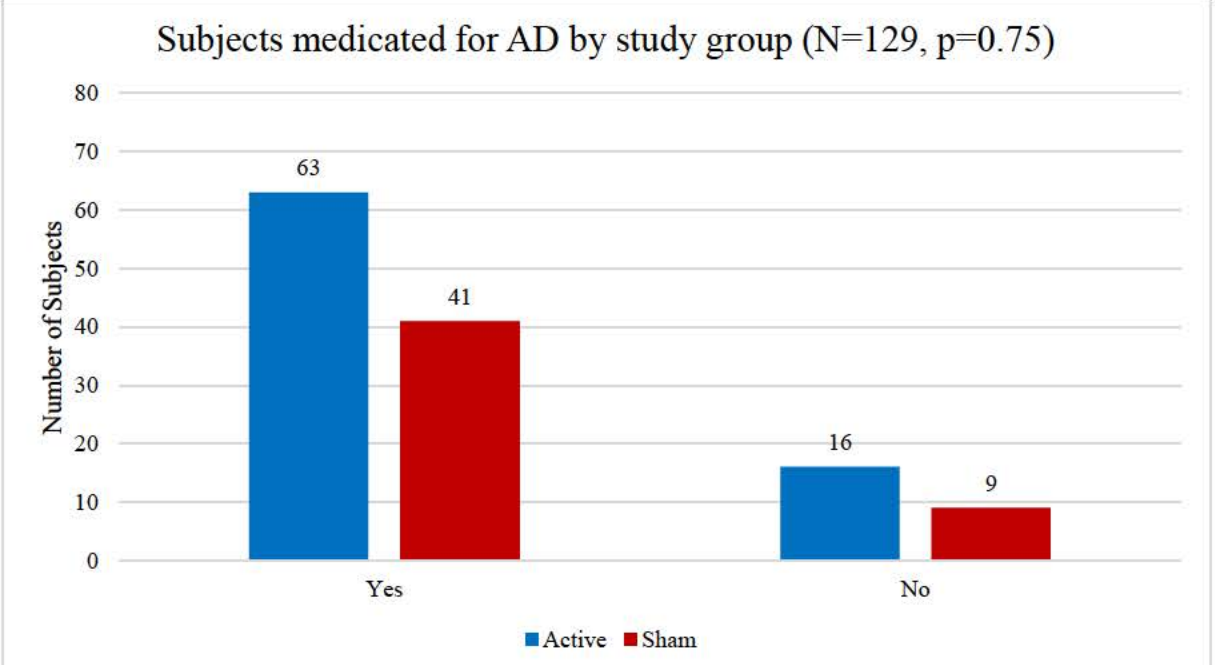
Table 10. Select Pivotal Study Baseline Characteristics

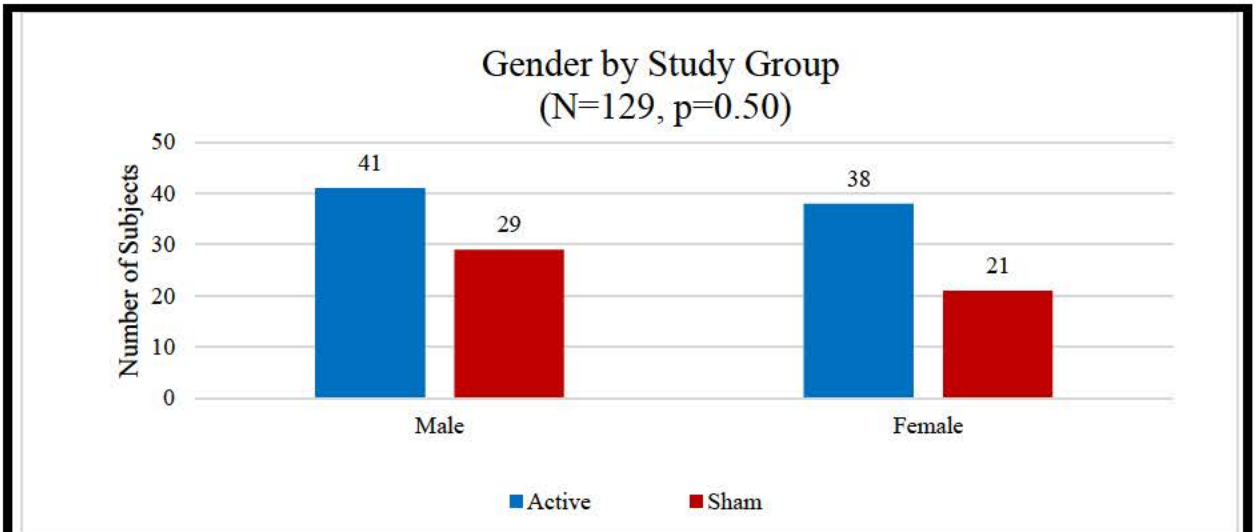


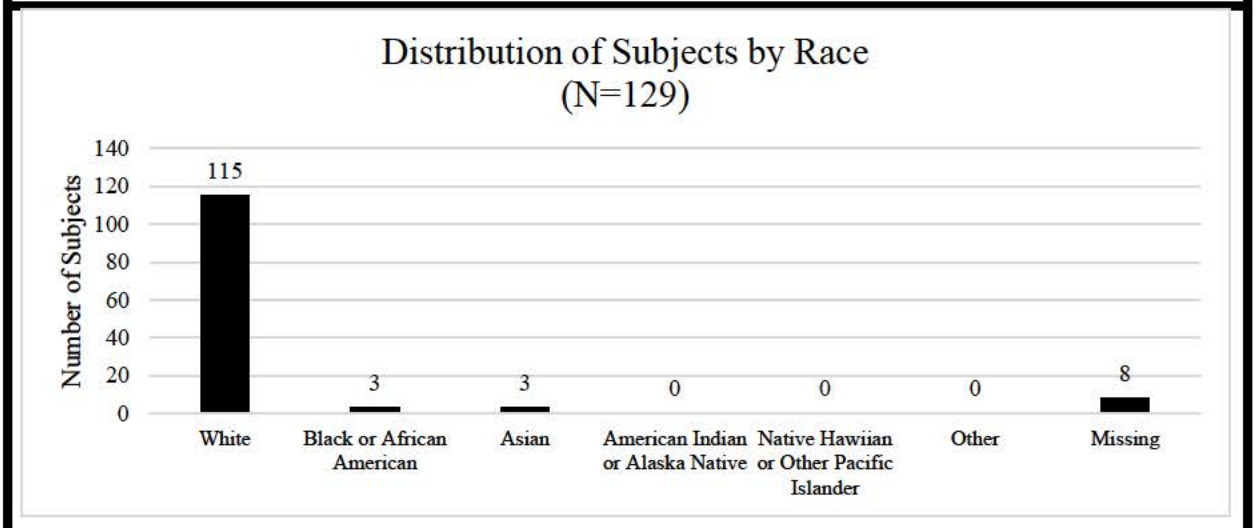


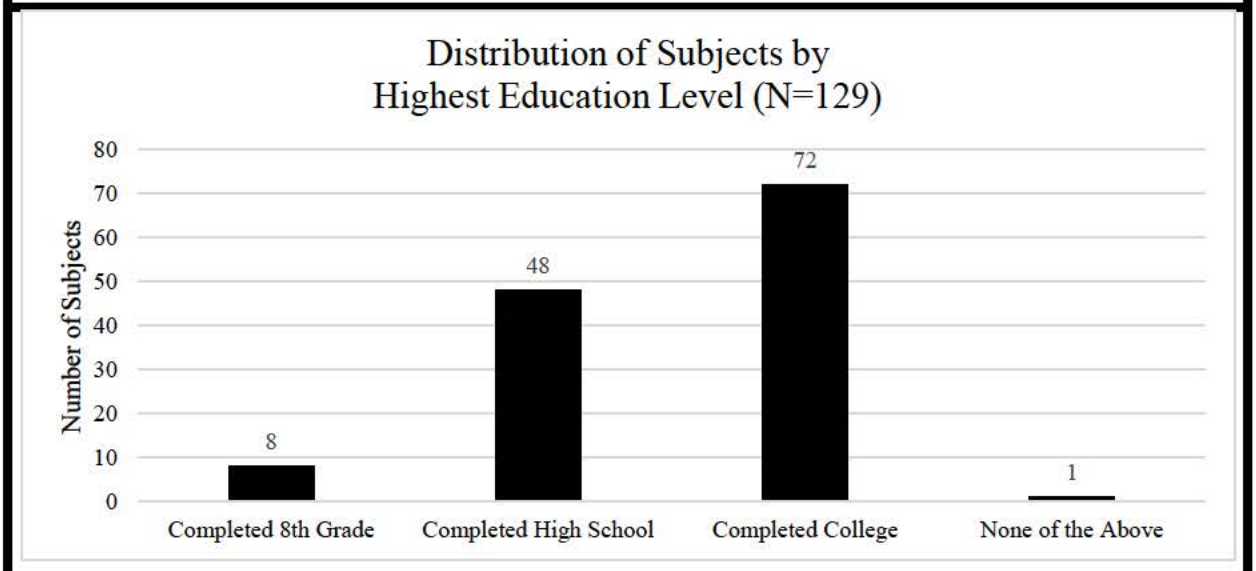
*Shaded bins (i.e., Baseline ADAS-Cog >30) indicate subjects who would not fall within the proposed indications for use.

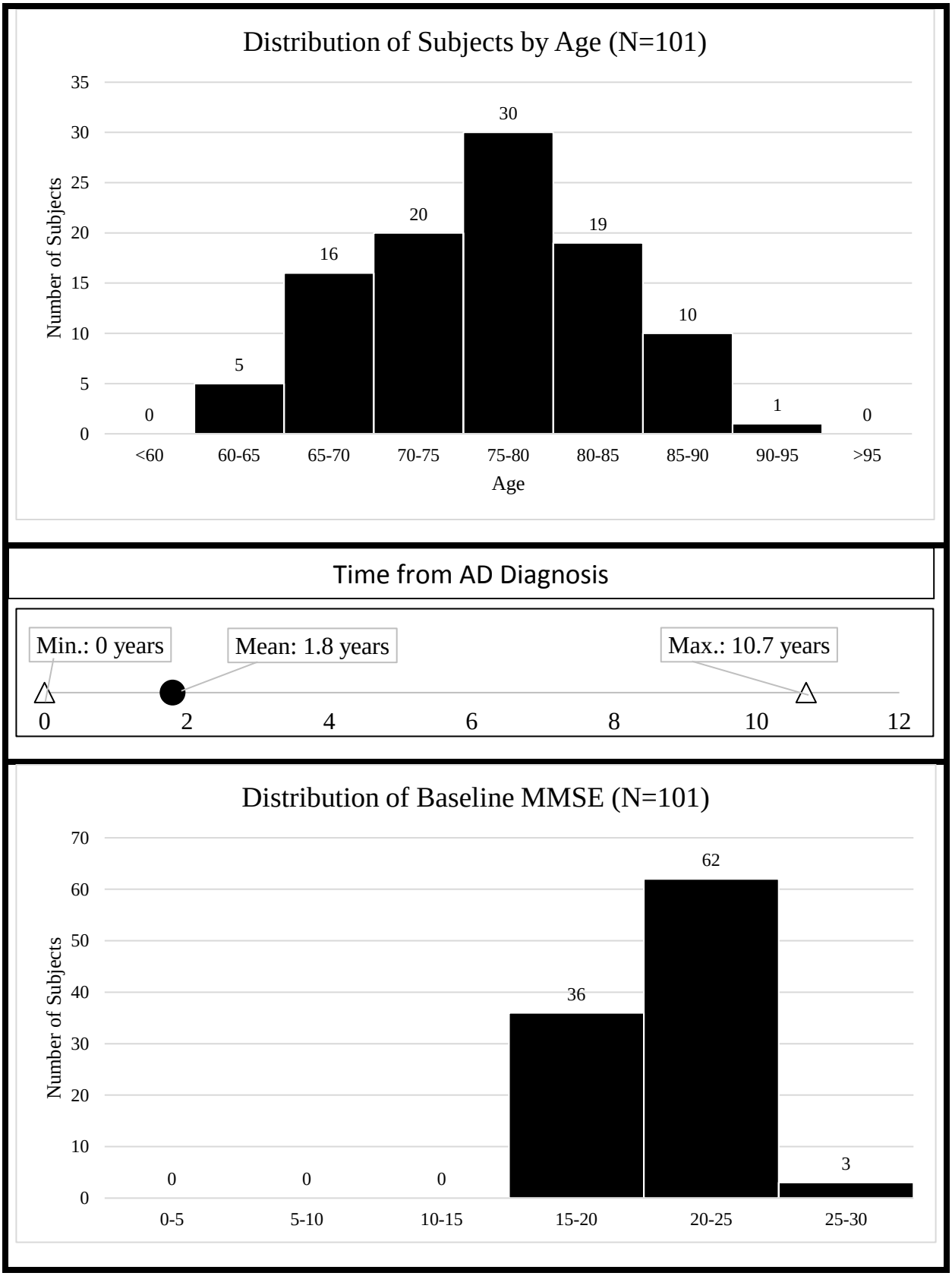
**Eligibility criteria restricted entry into the study to those with Baseline ADAS-Cog>17; subjects within bin 15-20 were all above 17.

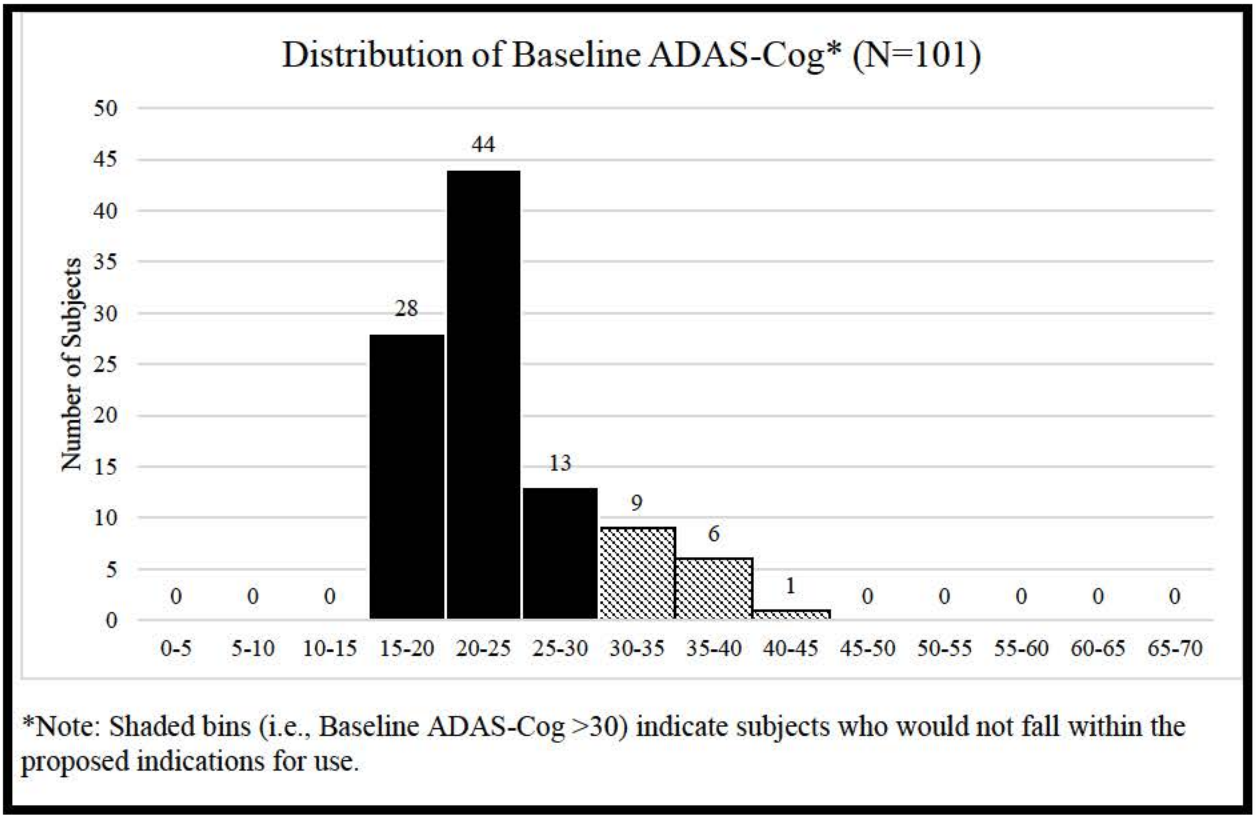



Gender by Study Group
(N=129, p=0.50)
Number of Subjects
50
40
30
20
10
0
41
29
38
21
Male
Female
Active
Sham


Distribution of Subjects by Race
(N=129)
Number of Subjects
140
120
100
80
60
40
20
0
115
3
3
0
0
0
8
White
Black or African American
Asian
American Indian or Alaska Native
Native Hawiian or Other Pacific Islander
Other
Missing


Distribution of Subjects by
Highest Education Level (N=129)
Number of Subjects
80
70
60
50
40
30
20
10
0
8
48
72
1
Completed 8th Grade
Completed High School
Completed College
None of the Above


Distribution of Subjects by Age (N=101)
Number of Subjects
35
30
25
20
15
10
5
0
0
5
16
20
30
19
10
1
0
<60
60-65
65-70
70-75
75-80
80-85
85-90
90-95
>95
Age
Time from AD Diagnosis
Min.: 0 years
Mean: 1.8 years
Max.: 10.7 years
0
2
4
6
8
10
12
Distribution of Baseline MMSE (N=101)
Number of Subjects
70
60
50
40
30
20
10
0
0
0
0
36
62
3
0-5
5-10
10-15
15-20
20-25
25-30



*Note: Shaded bins (i.e., Baseline ADAS-Cog >30) indicate subjects who would not fall within the proposed indications for use.

Protocol Deviations

There were several protocol deviations documented during the study as listed below.

Table 11. Pivotal Study - Protocol Deviations

Deviation Type	Number	Excluded from Analyses?	Comments
Did not meet inclusion criterion #4 – ADAS-Cog above 17	1	Yes	Withdrawn from study
Did not meet inclusion criterion #11 – Stable dose of AD drugs	1	Yes	Excluded from efficacy analysis
Different ADCS-CGI-C raters at baseline, week 7, and week 12	3	No	
Inconsistency in test order: Screening and Baseline procedures performed on same day; ADAS-Cog test performed after entire Screening process*	15	No	

Deviation Type	Number	Excluded from Analyses?	Comments
Different ADAS-Cog raters at baseline, week 7, and week 12	4	No	
ADAS-Cog rater had access to randomization information for subject but indicated no knowledge of assignment at time of evaluation	1	No	
ICF was not signed by the caregiver accompanying the enrolled subject	1	No	Resolved during study
Subject wrote on the ICF form but did not sign ICF form	1	No	Resolved during study
Consenting investigator failed to document that prospective study subject demonstrated the capacity to consent	2	No	Resolved during study

***Note:** The fifteen subjects that had the Screening and Baseline procedures performed on the same day, with the ADAS-Cog test performed following the entire Screening process were not excluded from the analyses. Per the Neuronix clinical study report submitted in DEN160053, this inconsistency in order of the tests may impact ADAS-Cog score results at Baseline. The sponsor states that analysis was performed with these patients excluded as well and it did not impact outcomes; however, FDA has not seen this analysis.

Follow-up Compliance

Out of the 109 randomized subjects 40.7% (n=24/59) of the Active group and 52.0% (n=26/50) of the Sham group completed the full series of 30 treatment sessions. 84.7% (n=50/59) of the Active group and 96.0% (n=48/60) of the Sham group completed at least 28 treatment sessions.

Out of 129 subjects included in the safety population, 8 subjects did not meet the minimum number of treatment visits as defined by the study protocol required to be eligible for inclusion in the Per Protocol population.

- Three subjects completed the 6-week treatment plan but did not participate in the minimum required treatment visits.
- Three subjects withdrew consent and did not complete the six weeks treatment plan
- One subject was withdrawn from the study and did not complete the six weeks treatment plan per sponsor request for not meeting inclusion-exclusion criteria
- One subject died during the course of the six weeks treatment plan (unrelated SAE subject 107-012 was found deceased after complaining of stomach pain)

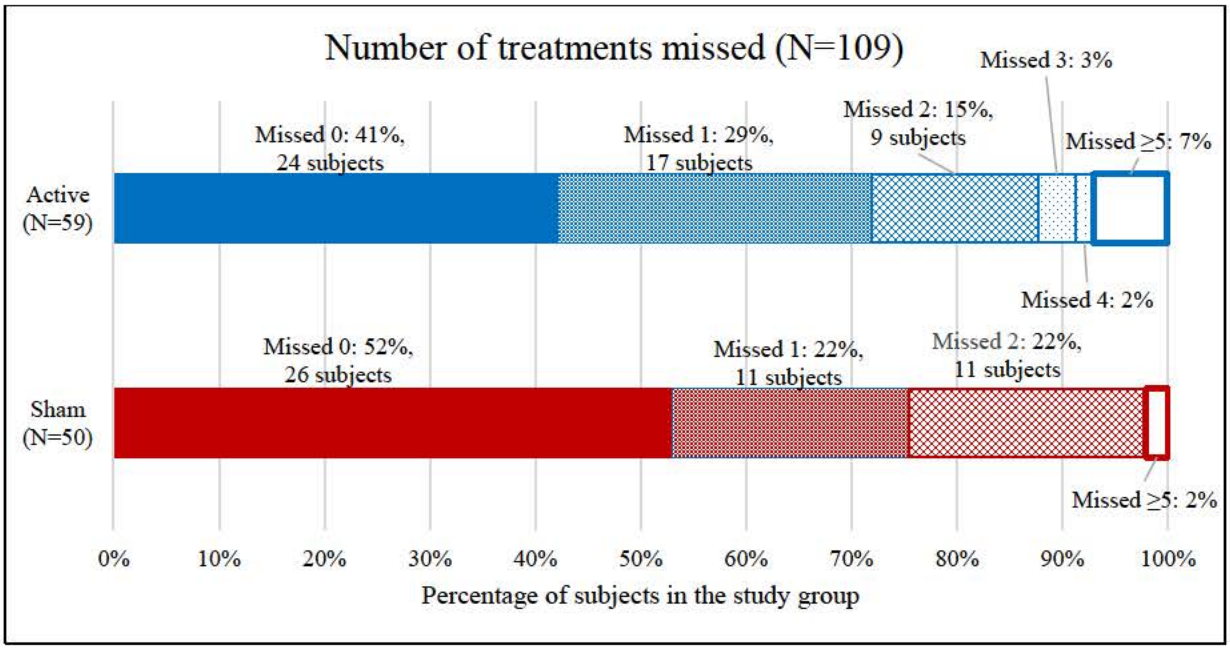


Figure 14. Pivotal Study - Subject Compliance

Safety Results

The safety population included 129 subjects. Overall 94 Adverse Events (AEs) were reported in the study. Of them, 70 AEs were of mild severity, 22 of moderate severity, and one was severe. The distribution of the severity of the AEs was similar in both study groups, with the single AE rated as 'Severe' occurring to a subject in the Treatment Group (the unrelated death described below).

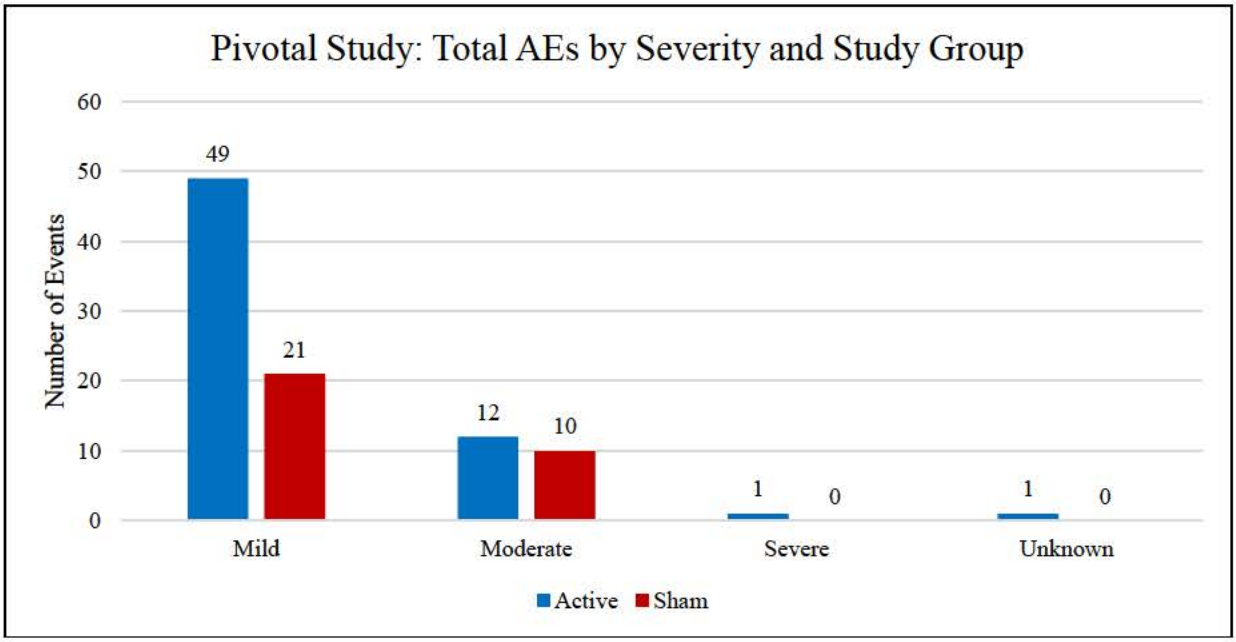


Figure 15. Pivotal Study - Total AEs by Severity and Study Group

AEs were also categorized by relationship to the Study Device. Overall, eleven subjects in the Treatment Group reported a total of 15 AEs which were found to be possibly, probably or definitely related to the investigational device. Two subjects in the Sham group also reported 4 AEs that were possibly, probably, or definitely related to the study device.

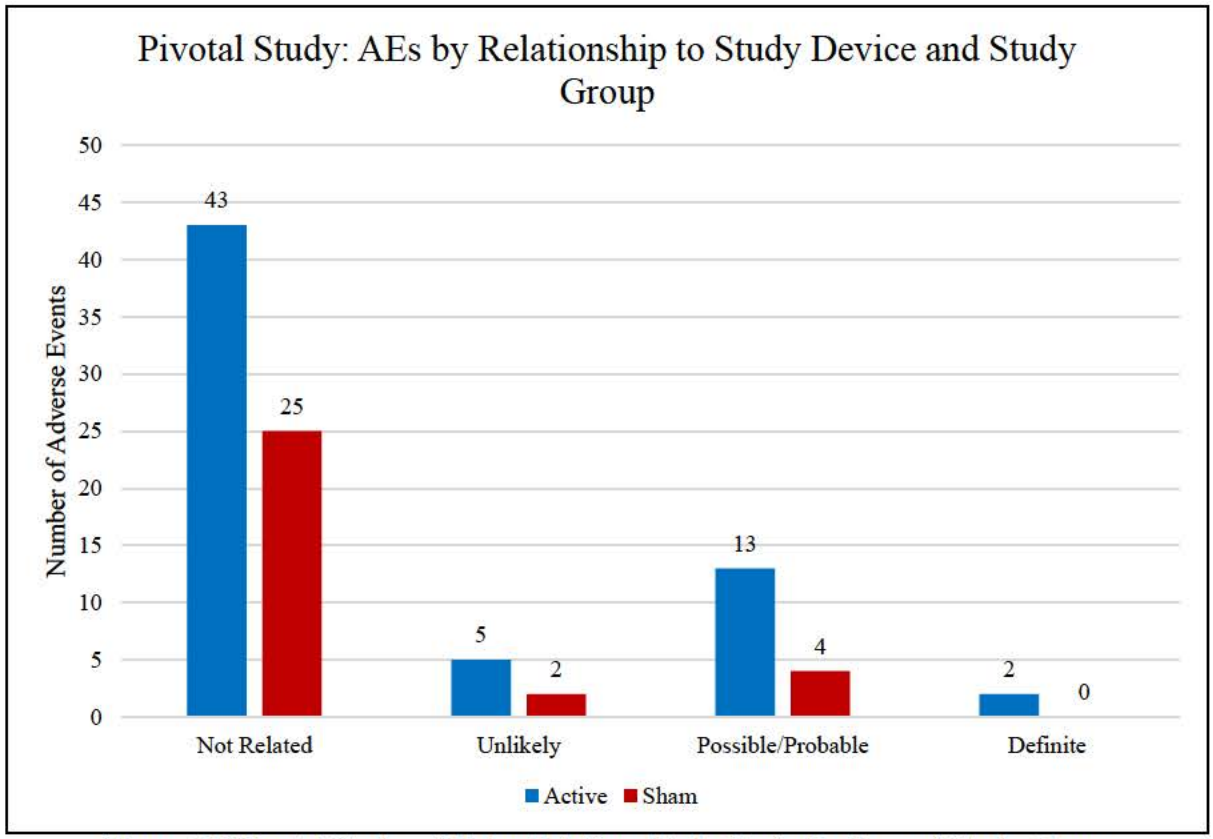


Figure 16. Pivotal Study - AEs by Relationship to Study Device and Study Group

Potentially related AEs were expected AEs that are often associated with TMS such as headache, neck pain, skin discomfort or muscle twitching. For some subjects, events persisted through multiple treatment sessions but severity was mild, did not require discontinuation, and the events were managed and overcome by adjusting/decreasing the treatment intensity (MT%) or administering Tylenol. These AEs were all transient, occurring during treatment and with no further side effect or other impact on subjects' daily life.

These 19 possible/probable/definite AEs are shown in Figure 17 below

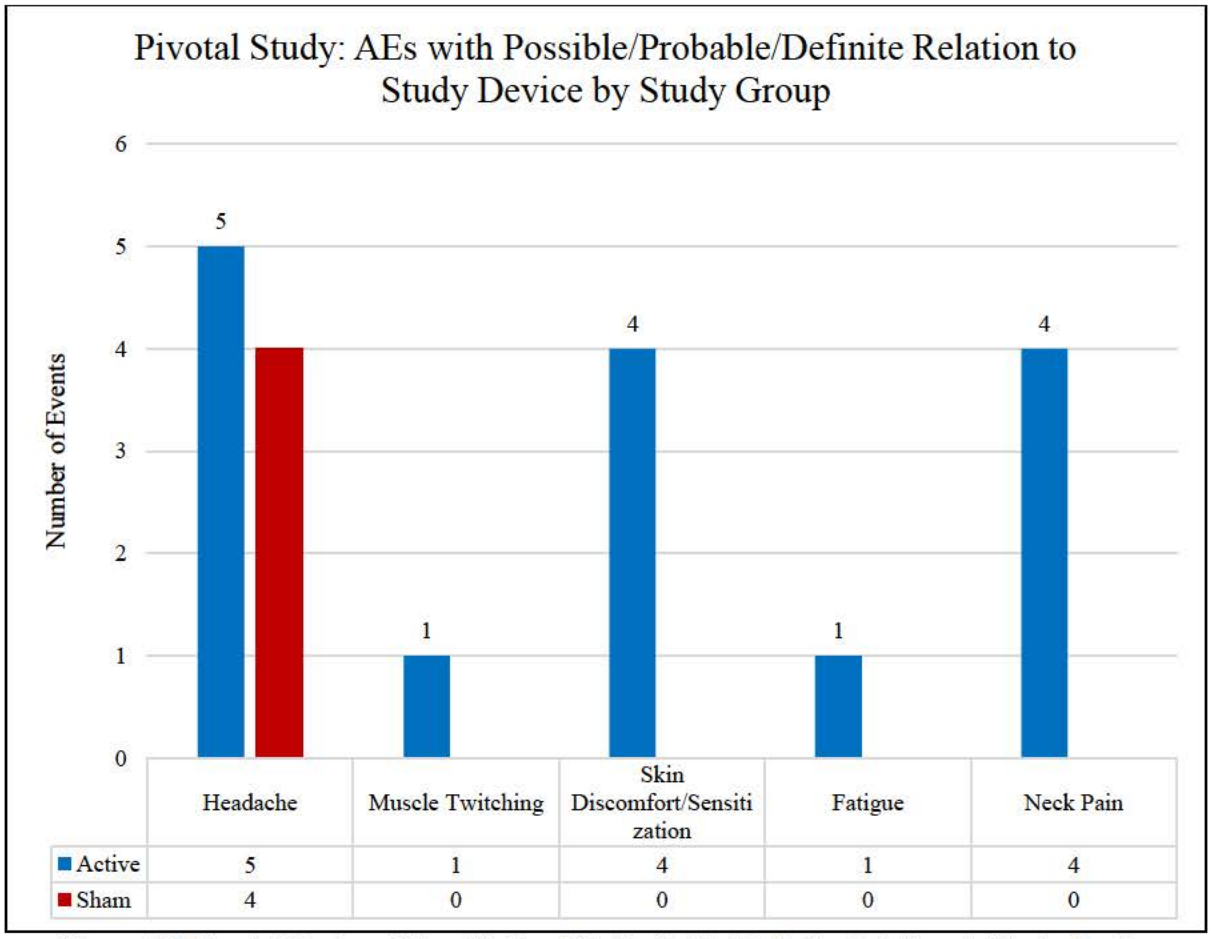


Figure 17. Pivotal Study - AEs with Possible/Probable/Definite Relation to Study Device

One death occurred in the study, which was unrelated to the study procedure or treatment. A description of this event is below:

- Subject (b) (6), Death: 83 year old female, randomized into the Treatment Group. The SAE occurred during the six week treatment course. Subject started treatments on October 13th, 2014. On November 3rd, 2014, the subject's caregiver contacted the site, stating the subject would miss her scheduled visit due to subject feeling tired. The following day, on November 4th, 2014, the subject's caregiver contacted the site again, reporting the subject had stomach issues on November 3rd, met with a physician and was prescribed with enema. Later that day, on the evening of November 3rd , the subject was found deceased in the bathroom. Autopsy was not performed. The last study treatment visit took place on Thursday, October 30th, 2014. The treatment session scheduled for Friday, October 31st, 2014 was cancelled in advance due to subject and caregiver travelling arrangements for the weekend. The event was assessed as unexpected SAE, not related to the study, by the site investigator, medical committee and sponsor. All study sites' IRBs were informed with the details of this event, as well as the FDA.

Three other serious adverse events occurred during the study. These were all determined unrelated to the study device or study procedures. These were not categorized by the sponsor as "severe" adverse events (see Figure 15) for which there is only one (the death) listed.

- Subject (b) (6) Cervical Fracture: 85 year old female, in the non-randomized Treatment Group. The SAE occurred post screening & baseline evaluations and prior to randomization and first treatment session. While scheduled to be randomized into the study, on March 5th, 2015, the subject experienced an upper respiratory infection and was referred by her primary care provider to inpatient hospitalization support. As the subject was preparing to leave her home to go to the hospital, she accidentally tripped and fell, injuring her neck. Injury resulted in a non-displaced fracture of the cervical spine that was treated with a hard cervical collar. Subject was hospitalized. The subject's participation in the study was discontinued. The event was resolved on June 25th, 2015. Subject did not receive any active or sham treatment.

- Subject (b) (6) Urinary Retention: 87 year old male, randomized into the Sham Group. Subject had history of BPH and urinary retention. SAE occurred during the six weeks treatment course. Subject started treatments on October 22th, 2014. On November 19th, 2014, the subject was presented to the hospital with suprapubic burning and pressure. A Foley catheter was placed. Subject was found to have UA positive for just nitrites. Given Ceftriaxone and admitted overnight. Event resolved on November 20th. Subject missed one study visit on November 20th, 2014, due to hospitalization.

- Subject (b) (6), Asthenia: 83 year old male, randomized into the Treatment Group. Subject had history of coronary artery bypass graft since July 2014. SAE occurred during the six weeks treatment course. Subject started treatments on April 20th, 2015. On April 24th, 2015, the subject arrived at study session number 5 disheveled and slow to respond. The site team discussed with the caregiver, and the subject was taken to the emergency room. No study procedures were performed on April 24th, 2015. Subject was hospitalized for generalized weakness. Subject found to be in rapid atrial fibrillation, have slightly low potassium, and mildly dehydrated. Normal EEG, CXR with mild atelectasis, cardiomegaly, head CT with no new findings. Subject was monitored, rehydrated, and discharged on April 30th, 2015. Subject missed four study sessions on April 24-30, 2015, due to hospitalization. Subject resumed study sessions on April 31st, 2015.

The most significant risk of TMS reported in the literature is inducement of seizures. No seizures were reported in this study.

Primary Effectiveness Results

As specified in the SAP, missing data for the primary and secondary endpoints were imputed using multiple imputation methodology. Because there was minimal missing data, the imputation was carried out and yielded results very similar to those obtained based on the non-imputed data; hence, the results described in their report were based on observed (non-imputed) data only. Additional analyses were carried out based on observed data only as well.

The primary endpoint was a change in the ADAS-Cog at seven weeks (one week following the end of treatment), as compared to the baseline measurement. The results of the pivotal study primary endpoint are shown graphically in Figure 18. The results show a difference of +1.45 points between groups, favoring the sham arm.

Table 12. Pivotal Study Primary Effectiveness Results

Analysis Population	Study Group	Change in ADAS-Cog from Baseline to Week 7						
		Mean	Std	Min	Median	Max	N	P-Value
Primary Efficacy	Treatment Group	0.07	3.97	-9.00	-0.33	14.00	53	
	Sham Group	-1.38	4.62	-16.67	-1.17	10.67	48	
	Difference	+1.45						0.09

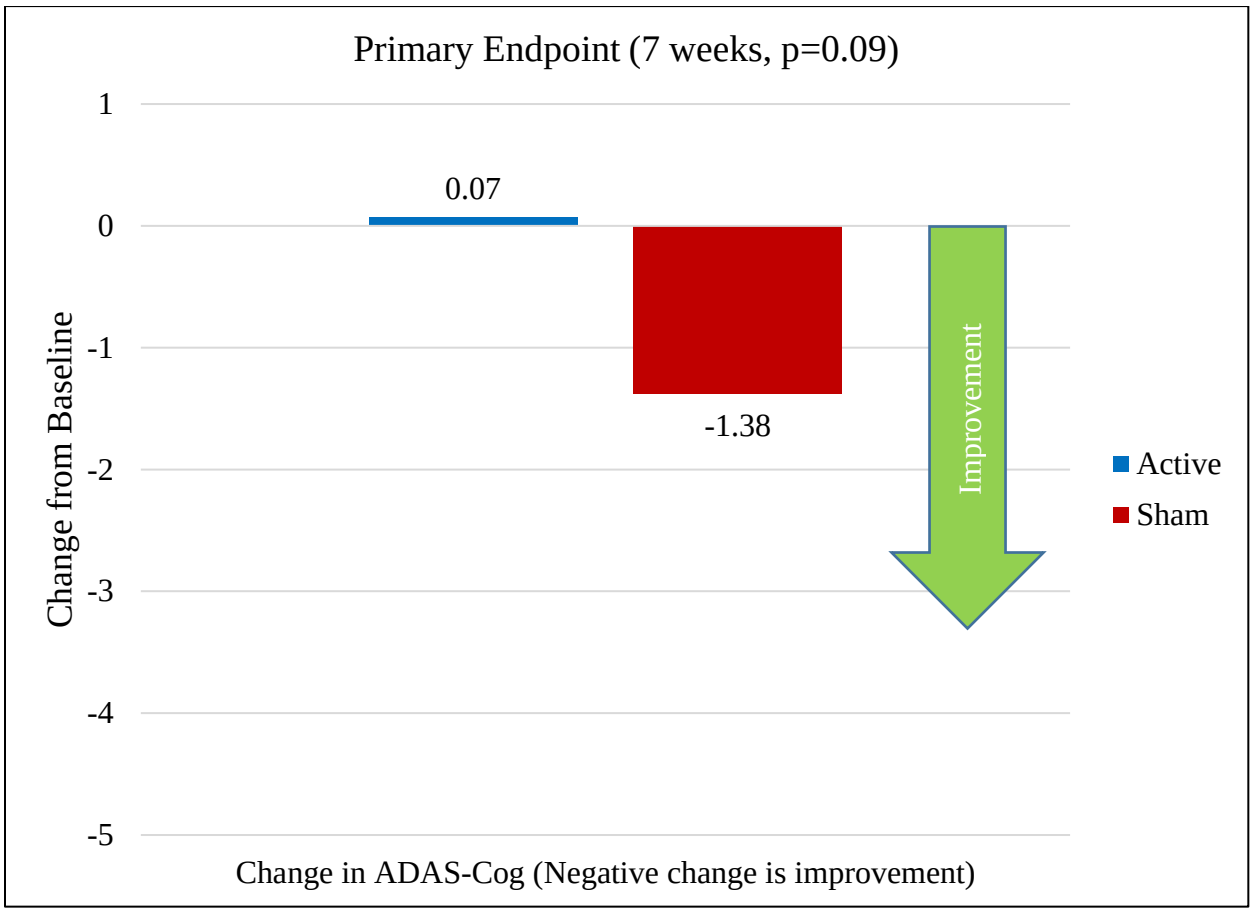


Figure 18. Pivotal Study Primary Endpoint Results, Group Changes from Baseline to 7 Weeks

To understand the magnitude of these changes from baseline with respect to the entire ADAS-Cog scale, the FDA provides Figure 19 below. This figure provides the mean baseline ADAS-Cog values for the active and sham groups out of the full 70-point scale and shows the full ADAS-Cog scores for each group at the 7 week primary endpoint.

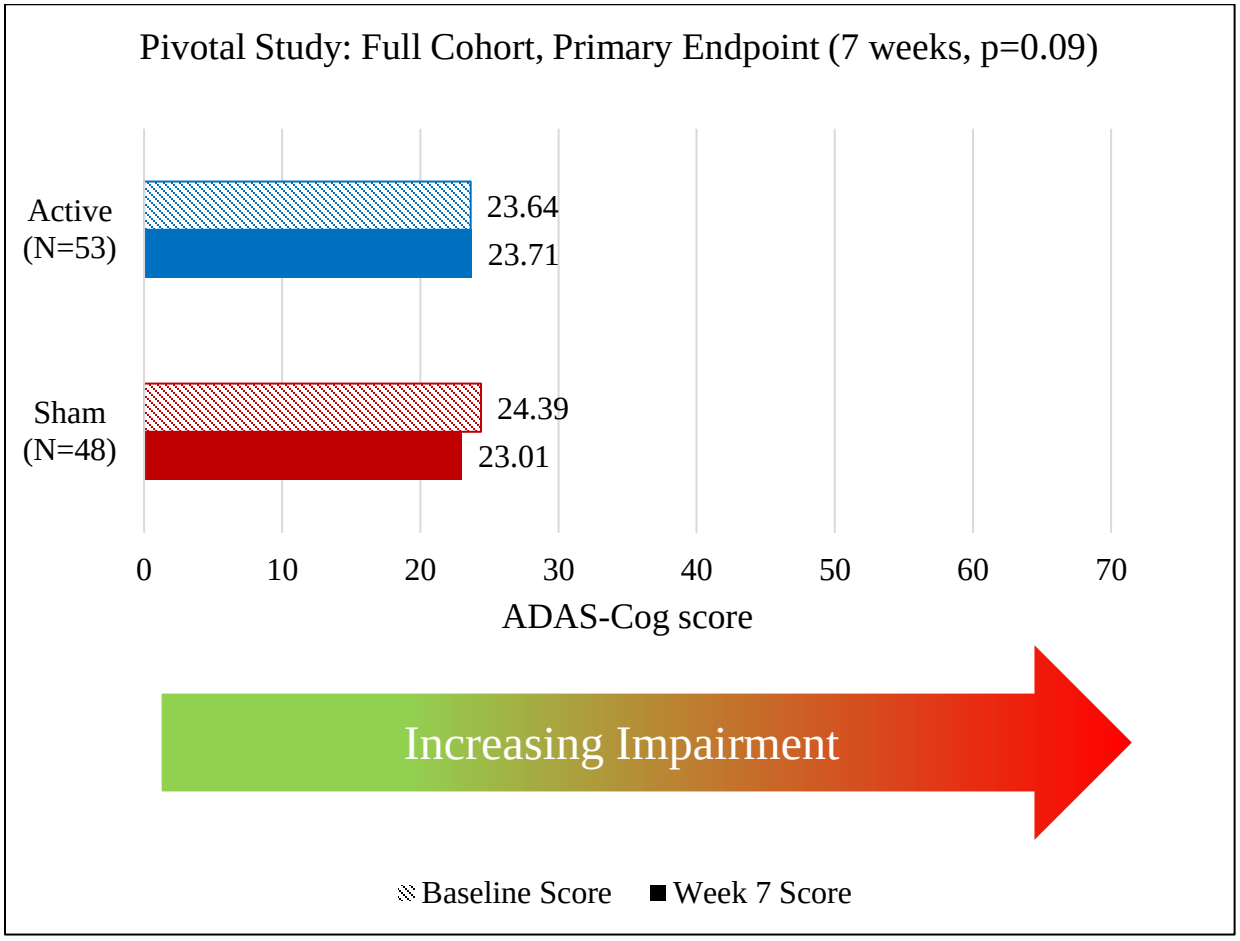


Figure 19. Pivotal Study Primary Endpoint Results, Group Means at Baseline and 7 Weeks

Figure 20 provides the individual ADAS-Cog results for active and sham subjects from baseline to the 7-week primary endpoint timepoint. In this view, it is possible to assess how individual subjects contributed to the average worsening seen in the active group and the difference of +1.45 points (favoring sham) between groups.

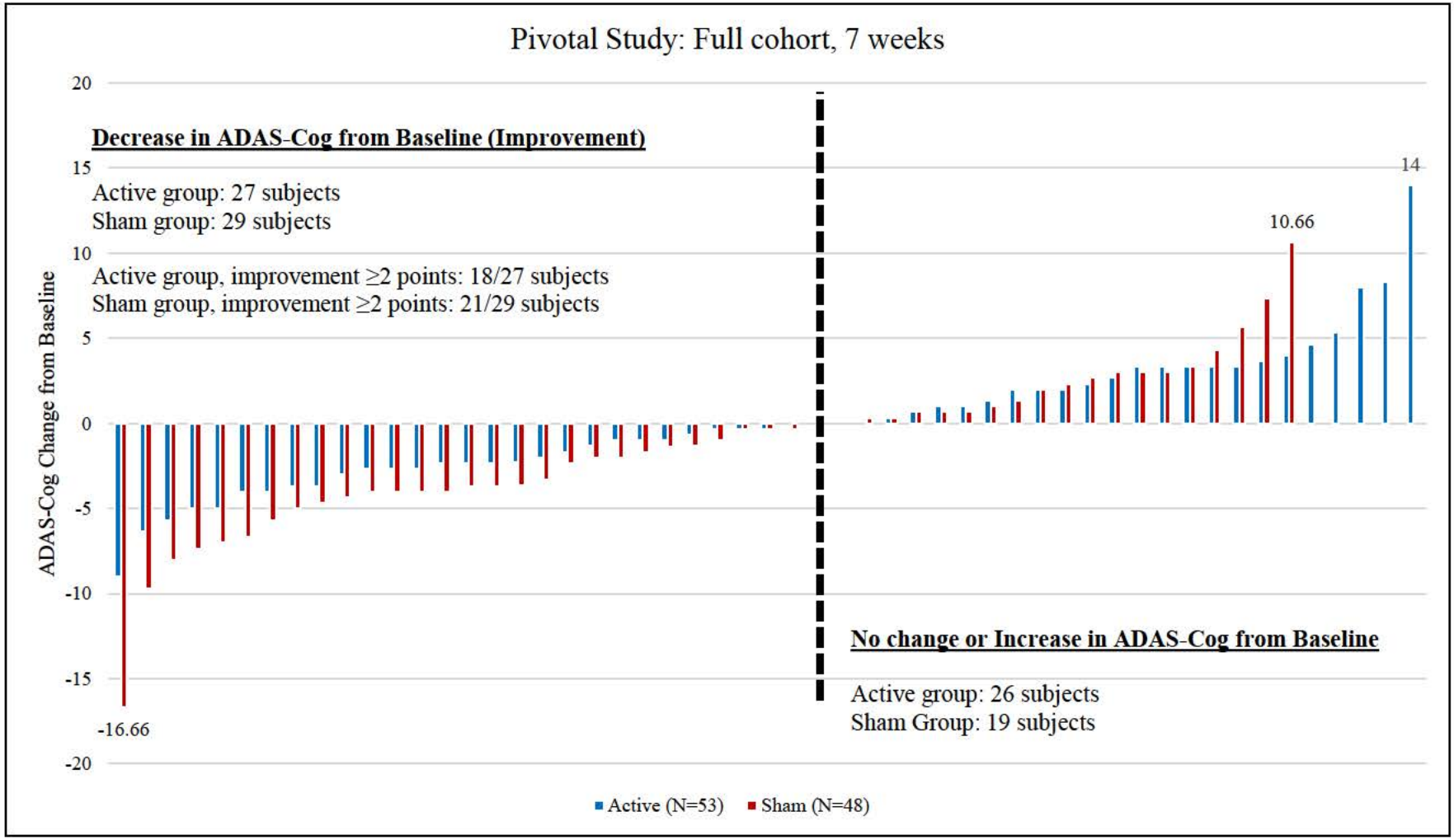


Figure 20. Pivotal Study Primary Endpoint Results, Individual ADAS-Cog Results, Active and Sham, 7 Weeks

Secondary Endpoint Results

The study design included three, pre-specified secondary endpoints as follows:

1) Change in ADCS-CGIC from Baseline to week 7
2) Change in ADCS-CGIC from Baseline to week 12 (6 weeks after discontinuation of treatment)
3) Change in ADAS-Cog from Baseline to week 12 (6 weeks after discontinuation of treatment)

Results of the secondary endpoints are displayed in the figures below:

Change in ADCS-CGIC from Baseline to week 7

Figure 21 provides a line graph showing the reported CGI-C mean scores of each group at 7wks. As noted in the Clinical Context section above, a score of "4" on the CGI-C indicates no clinical change. The results indicate a mean difference between groups of 0.02 in favor of treatment (p=0.96).

Figure 21. Pivotal Study Secondary Endpoint, CGI-C at 7wks, Group Means (p=0.96)

Figure 22 provides a breakdown of the individual CGI-C results for the entire CGI-scale for both active and sham groups at this 7wk timepoint. A score of "1" would indicate "very much improved" while a score of "7" would indicate "very much worse." The associated p-value comes from a comparison of the two treatment groups on the distribution of CGI-C scores across the 1-7 range.

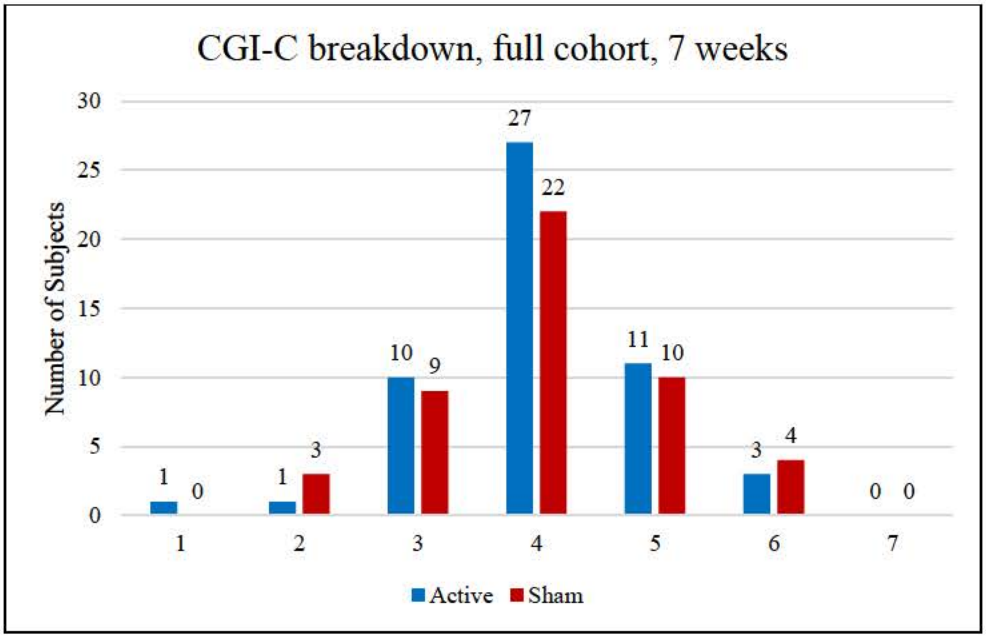


Figure 22. Pivotal Study Secondary Endpoint, CGI-C at 7wks, Breakdown (p=0.78)

Change in ADCS-CGIC from Baseline to week 12

Figure 23 shows the same CGI-C full scale for the 12-week endpoint results. Please note that this is approximately 6 weeks after the last treatment visit. The results indicate a mean difference between groups of 0.35 in favor of treatment (p=0.12).

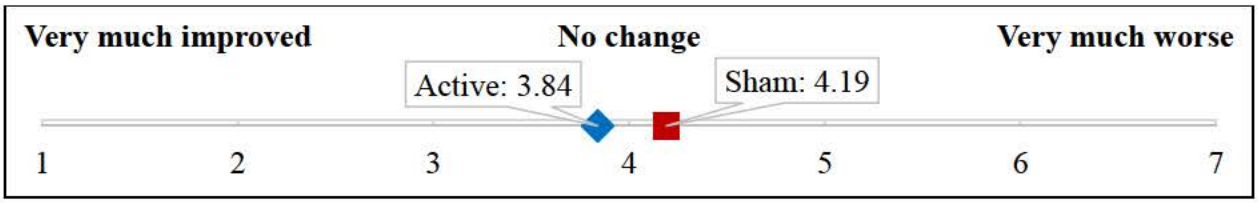


Figure 23. Pivotal Study Secondary Endpoint, CGI-C at 12wks, Group Means (p=0.12)

Figure 24 provides a breakdown of the individual CGI-C results for the entire CGI-scale for both active and sham groups at this 12wk timepoint. Again, a score of "1" would indicate "very much improved" while a score of "7" would indicate "very much worse". The associated p-value comes from a comparison of the two treatment groups on the distribution of CGI-C scores across the 1-7 range. From Figure 24, it appears that the greatest difference between Active and Sham is in category 4 ("no change") versus 5 ("worsening").

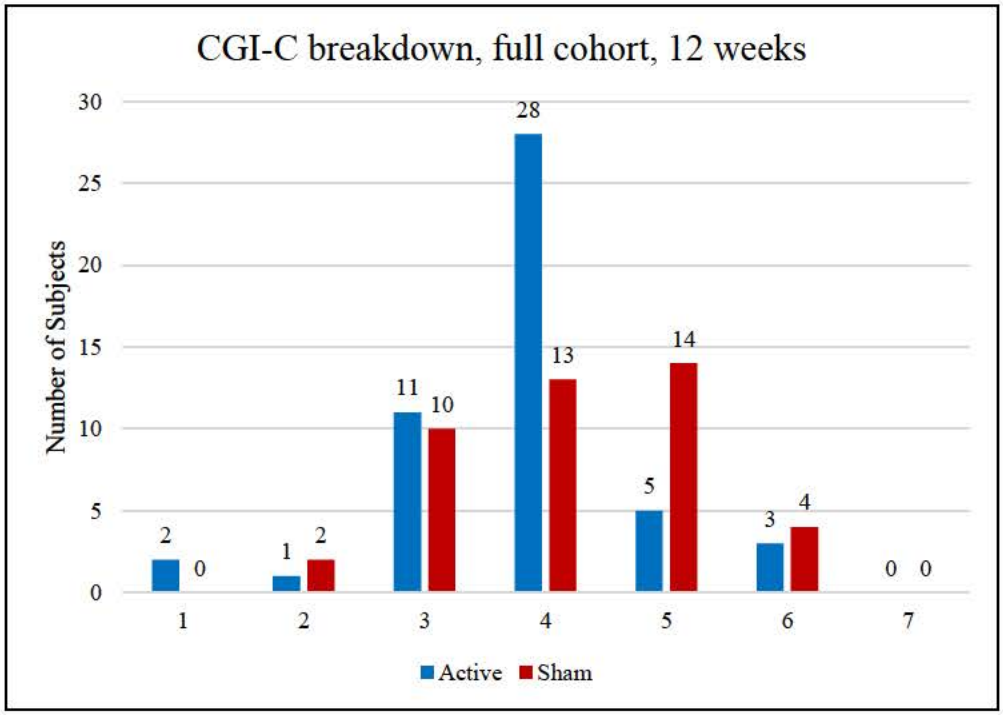


Figure 24. Pivotal Study Secondary Endpoint, CGI-C at 12wks, Breakdown (p=0.04)

Change in ADAS-Cog from Baseline to week 12

Neuronix also collected data from ADAS-Cog assessments conducted at 12 weeks post treatment (approximately 6 weeks post final treatment visit). Please note that there was no neuroAD active or sham therapy provided between the assessments conducted at 7 weeks and those conducted at 12 weeks. Figure 25 shows the mean of the active and sham groups on the change in ADAS-Cog from baseline to 12 weeks. The results indicate a difference between groups of -0.42 in favor of treatment.

Table 13. Pivotal Study Secondary Effectiveness Endpoint, ADAS-Cog at 12 weeks

Analysis Population	Study Group	Change in ADAS-Cog from Baseline to Week 12						
		Mean	Std	Min	Median	Max	N	P-Value
Primary Efficacy Analysis	Treatment Group	-1.03	4.85	-11.67	-1.33	10.67	51	
	Sham Group	-0.61	3.96	-10.67	-1.00	8.00	47	
	Difference	-0.42						0.64

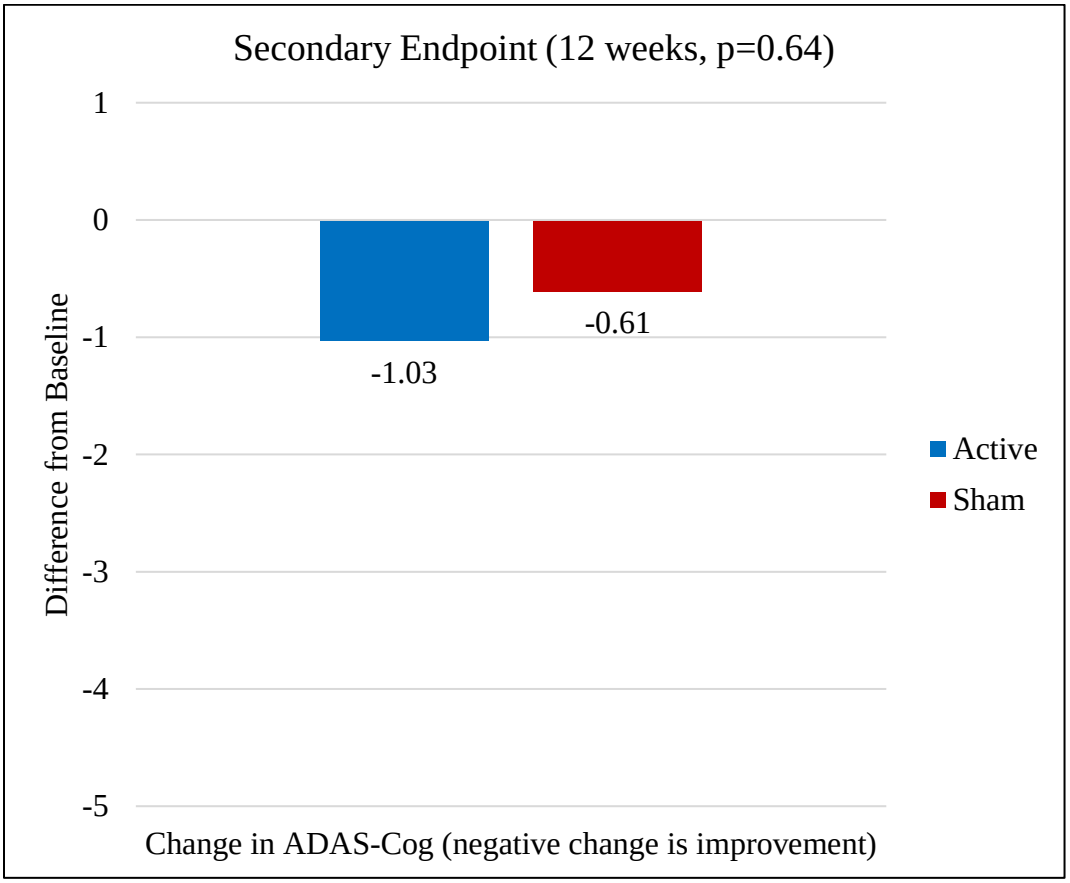


Figure 25. Pivotal Study Secondary Endpoint Results, ADAS-Cog Group Changes from Baseline to 12wks (p=0.64)

As for the primary endpoint, Figure 26 below represents the magnitude of these changes from baseline with respect to the entire ADAS-Cog scale. This figure provides the mean baseline ADAS-Cog values for the active and sham groups out of the full 70-point scale and shows the full ADAS-Cog scores for each group at the 12 week secondary endpoint.

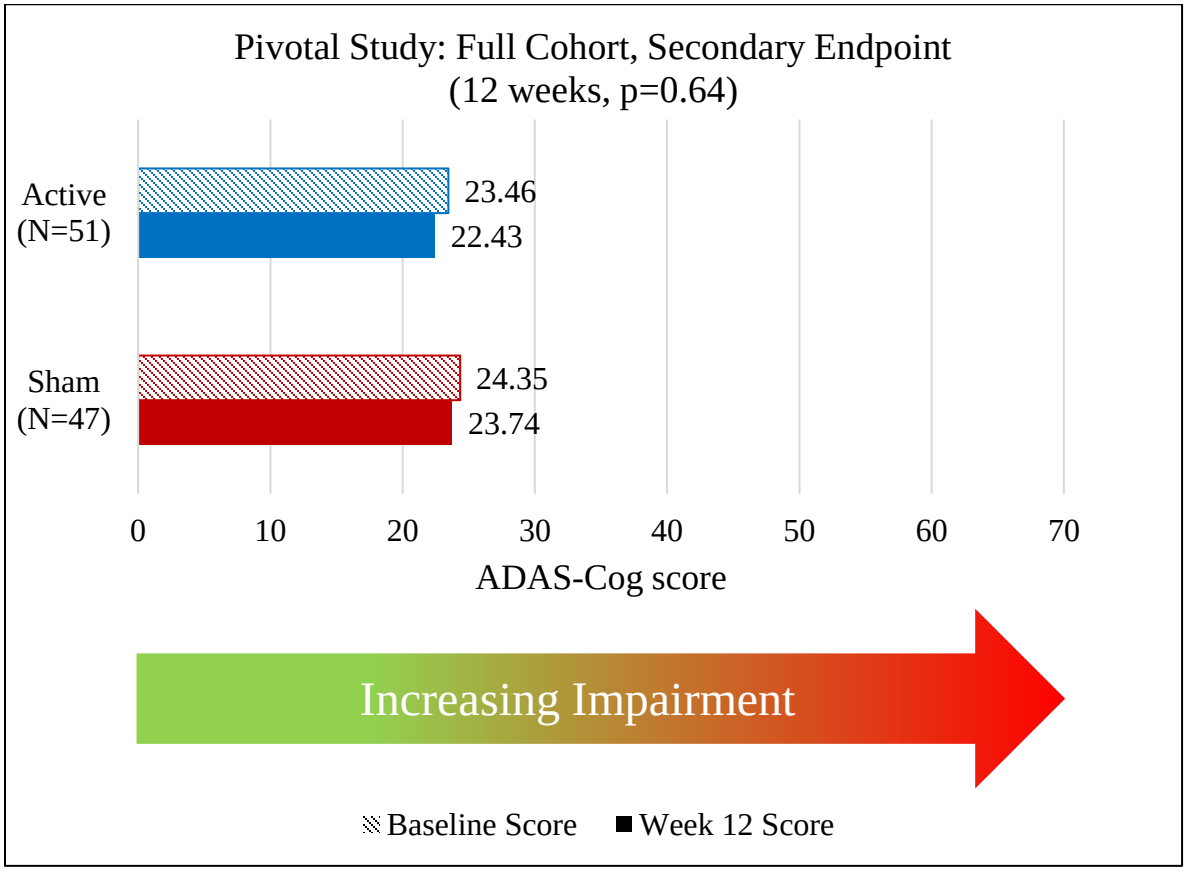


Figure 26. Pivotal Study Secondary Endpoint Results, Group Means at Baseline and 12 Weeks

As for the ADAS-Cog results at the primary endpoint assessed at 7 weeks, the figure below shows individual results for the active (blue) and sham (red) subjects to better understand the overall mean changes noted.

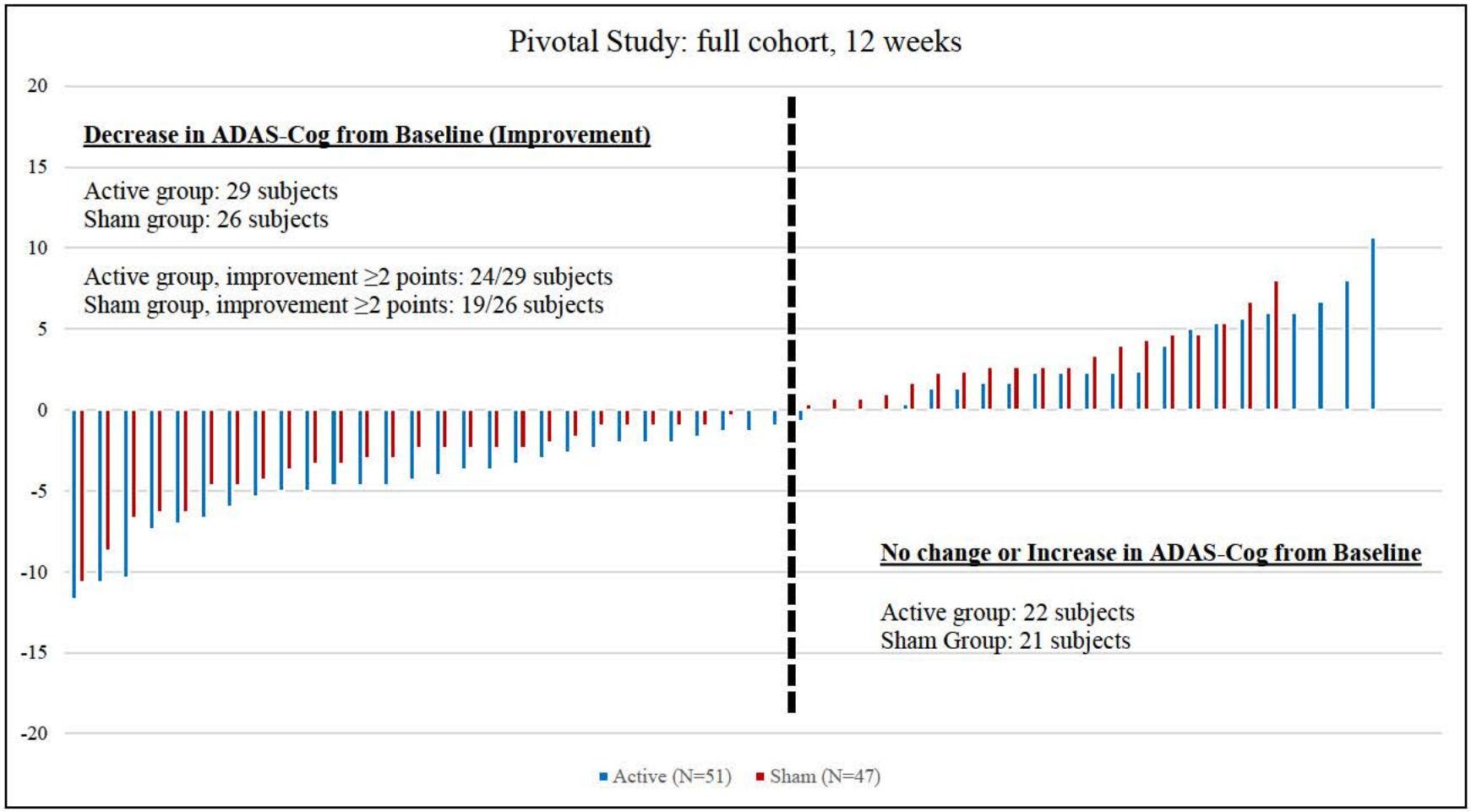


Figure 27. Pivotal Study - Individual ADAS-Cog Results, Active and Sham, 12 weeks

FDA Summary Comments on Pivotal Clinical Study

Safety

The clinical study report provided information on the overall number of events by treatment group categorized as mild, moderate, or severe. However, detailed adverse event information was limited only to those AEs which were determined to be related to the study device or procedure as assessed by the study site investigator; this detailed information is provided in Figure 17. Most of the AEs reported in the study were found to be not related or unlikely to be related to study procedures. All serious adverse events, including one death, were reported regardless of relationship to study device.

The rate of pivotal study subjects experiencing any definite, probable or possible study procedure or device-related adverse event was 14% (11/79, includes the 20 non-randomized active patients) in the active group and 4% (2/50) in the sham group. The 11 active group subjects reported 15 events total that were found to correlate with relatedness to the device or procedure. These AEs were reported and determined by the site investigator. The AEs were described to be mild per the clinical study report and were all among the expected AEs previously reported with TMS. The AEs included headache, neck pain, fatigue, skin discomfort, and muscle twitching. The clinical study report states that some of the AEs persisted through multiple treatment sessions which were described as mild, transient, not requiring treatment discontinuation, occurring during treatment, and were managed by adjusting the treatment intensity (MT%). All adverse events related to the study device resolved without sequalae.

While seizures are a known risk of TMS procedures, no seizures were noted in the neuroAD pivotal trial.

Based on the adverse events recorded during the neuroAD pivotal trial, the neuroAD appears to carry a higher risk than the sham device. However, the risk of the neuroAD device appears to be low.

Effectiveness

The summary figure below demonstrates the results of the group means (Active and Sham) for the ADAS-Cog results from baseline to 7 weeks (Primary Endpoint) and from baseline to 12 weeks (Secondary Endpoint). 53 active subjects and 48 sham subjects completed the 7-week visit, and 51 active subjects and 47 sham subjects completed the 12 week visit.

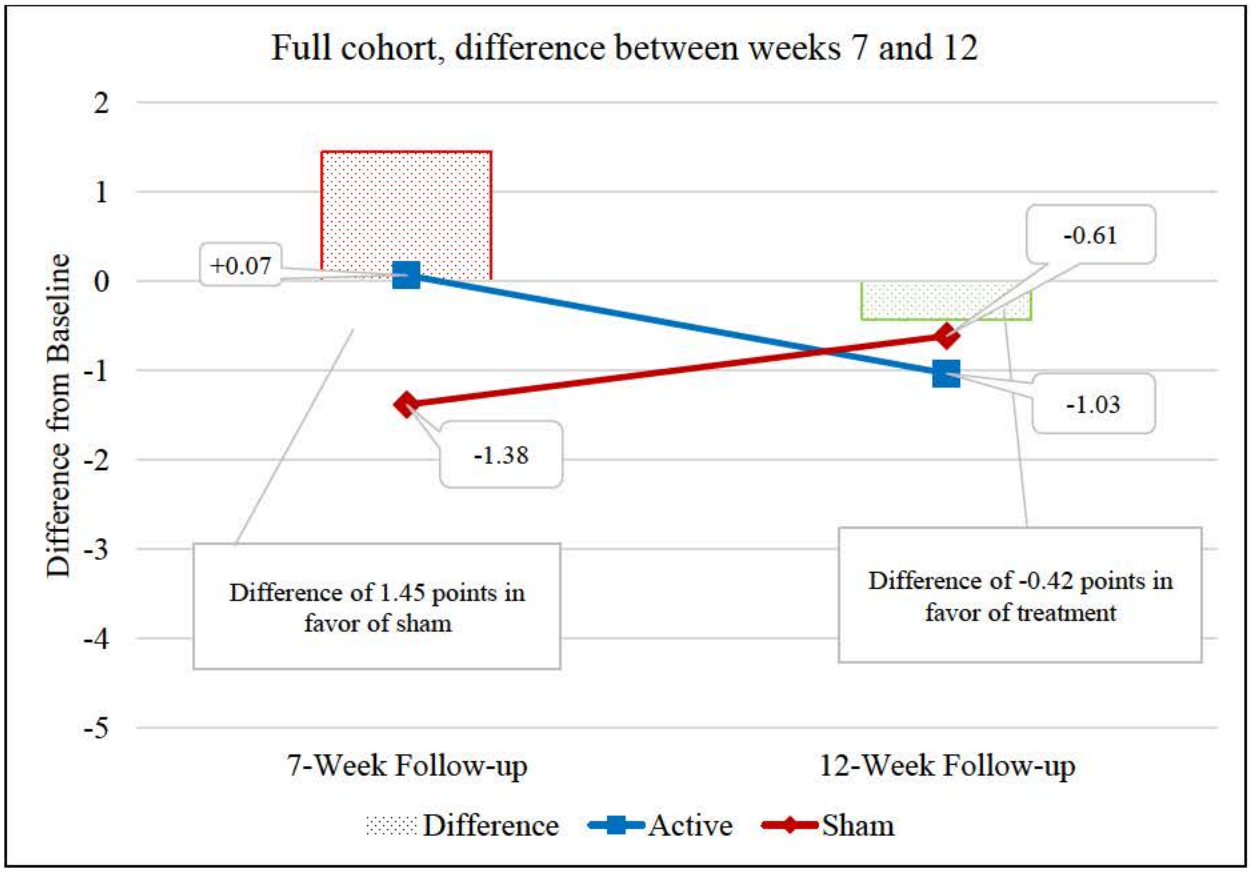


Figure 28. Pivotal Study - Mean Change in ADAS-Cog Summary Results

At the primary endpoint, the sham performs better than the active group by +1.45 points. In this case, the difference is caused by a numerical worsening from baseline in the active group (+0.07) and a numerical improvement in the sham group (-1.38). This does not provide evidence in support of device effectiveness.

Neuronix also analyzed the secondary endpoint of the ADAS-Cog at 12 weeks. At the secondary endpoint timepoint of 12 weeks (approximately 6 weeks post-final treatment) the group means for the ADAS-Cog results reverse direction from their 7-week scores. At 12 weeks the active group performs better on average than the sham group by -0.42 points. At 12 weeks the sham group still exhibits numerical improvement on average (-0.61), but by a smaller margin than at the 7-week endpoint (i.e., the sham group begins to trend back towards baseline from 7-12 weeks). The active group also shows numerical improvement (-1.03) for the first time in the study at 12 weeks. It appears that the active group improves after the treatment ceases for six weeks. This "recovery" in the active group between 7 and 12 weeks occurred in the absence of any neuroAD therapy.

FDA also looked to the secondary assessment scale, the CGI-C, to provide clinical context to the ADAS-Cog results. However, the clinical global impression of change based on CGI-C assessment at the 7-week assessment timepoint only indicated a difference between groups of 0.02 in favor of treatment. The results at week 12 (approximately 6 weeks after the last treatment visit) indicate a difference between groups of 0.35 in favor of treatment. It is not clear whether these changes lend any support to the clinical meaning of the changes found in the ADAS-Cog scale.

Based on the CGI-C results of the pivotal trial, published work, and expert clinical opinion, it is unclear that the most favorable result for the neuroAD of -0.42-points in favor of treatment at 12 weeks represents a clinically meaningful benefit (please see Benefit-Risk Assessment section of this memo).

This 12 week result must also be considered in light of the result of the primary endpoint, the effect seen at 12 weeks is in the absence of any neuroAD intervention given between 7 weeks and 12 weeks, and this trend in results shown in the pivotal study between 7 weeks and 12 weeks has not been observed in other studies (please see Appendix V. Post Hoc Analyses Using Supplemental Clinical Data Sources). Additionally, it is uncertain whether the result at 12 weeks should be attributed to the neuroAD intervention given that there is evidence that repeat assessments on cognitive testing result in measured improvement in results based on familiarity and practice effects (Heilbronner et al., 2010; Jacobs et al., 2017).Finally, the p-value at 12 weeks is of such large magnitude (0.64) that it would not support rejecting the hypothesis of no difference between Active and Sham, even if the assessment had been primary.

Panel Question: The Panel will be asked to discuss and make recommendations on whether the U.S. pivotal study demonstrates a clinically meaningful benefit for the neuroAD as an adjunctive therapy

Pivotal Clinical Study – Post-Hoc Analysis of Baseline ADAS-Cog≤30 Subgroup

In addition to the pre-specified analysis of the pivotal study that is discussed above, in the original submission Neuronix also presented a post-hoc analysis using the pivotal study data with only those patients with a baseline ADAS-Cog score ≤30 (hereafter also termed the "subgroup" or the "indicated population). It is intended that the subgroup of subjects in the pivotal study that had a baseline ADAS-Cog score that was ≤ 30 define a clinically plausible subset of mild to moderate Alzheimer's disease patients. This subset eliminated 8 subjects from the treatment group and 8 subjects from the sham group with baseline ADAS-Cog scores > 30. As this subgroup is the foundation of the proposed patient population and is comprised of data from the pivotal study, this subgroup analysis is presented in the body of the Executive Summary.

The primary analysis given in the SAP included a test of interaction between treatment group and baseline ADAS-Cog, at 7 weeks. If the interaction test was not statistically significant, the interaction term would be eliminated from a statistical model. The SAP did not further specify the analysis in the case of a statistically significant interaction. Note that the analysis showed a statistically significant interaction between treatment group and baseline ADAS-Cog at 7 weeks. The observed interaction at the secondary time point of 12 weeks appeared to be even stronger.

Neuronix also provided several additional post-hoc analyses using the pivotal study data as well as supplemental datasets. These additional post-hoc analyses and supplemental investigations are discussed in more detail in the appendices.

For all post-hoc analyses, Neuronix presented data on the original primary assessment timepoint of 7 weeks as well as the secondary assessment timepoint of 12 weeks.

Note: The analyses presented in this section are post-hoc analyses without pre-specification and multiplicity adjustment. Therefore, we do not include p-values.

Post-Hoc Analysis Results – ADAS-Cog≤30 Subgroup

The figures below illustrate the post-hoc analysis subgroup of subjects in the pivotal study with a baseline ADAS-Cog≤30. Results are shown for both the ADAS-Cog and the CGI-C scales at both assessment timepoints of 7 and 12 weeks. There was no neuroAD active or sham treatment between the assessment conducted at 7 weeks and at 12 weeks.

As in the entire cohort, at the 7 week timepoint the ADAS-Cog mean scores in this subgroup continue to favor the sham group, but by a smaller margin of +0.47 points.

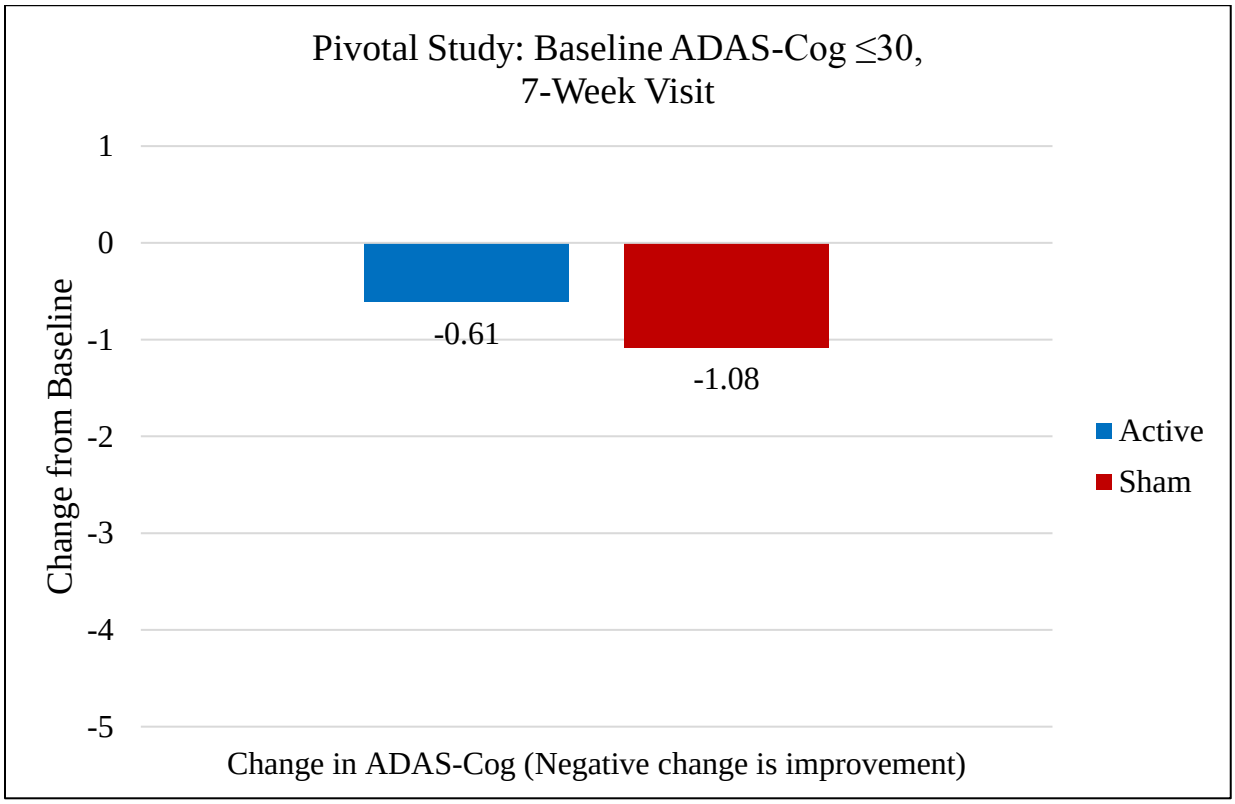


Figure 29. Pivotal Study Post-Hoc Analysis Results, Subgroup, ADAS-Cog Group Changes from Baseline at to 7wks

These results are shown in Figure 30 below as the group means on actual ADAS-Cog score based on the 0-70 points scale.

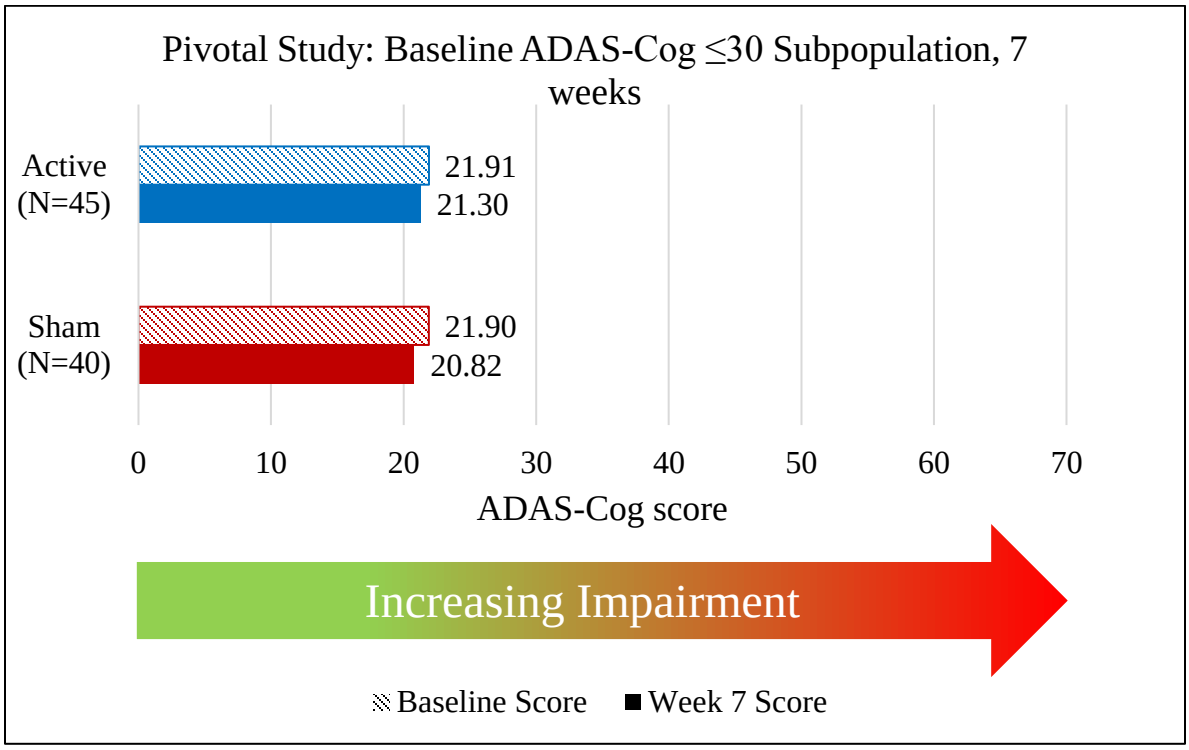


Figure 30. Pivotal Study Post-Hoc Analysis Results, Subgroup, ADAS-Cog Group Means at Baseline and 7 Weeks

At 12 weeks, this trend reverses and the difference between groups favors the treatment group by -1.61 points. This value of -1.61 was the largest difference in favor of treatment shown in the pivotal study dataset.

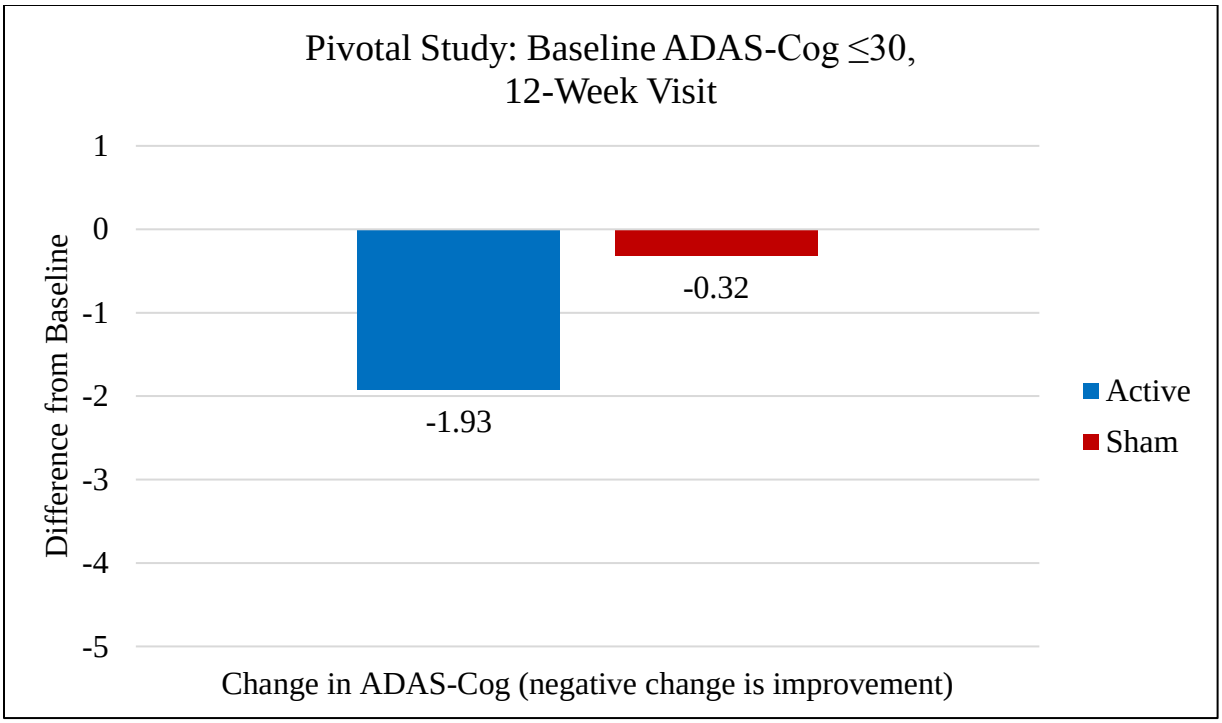


Figure 31. Pivotal Study Post-Hoc Analysis Results, Subgroup, ADAS-Cog Group Changes from Baseline to at 12wks

Again, we also present these results in Figure 32 below as the group means on actual ADAS-Cog score based on the 0-70 points scale.

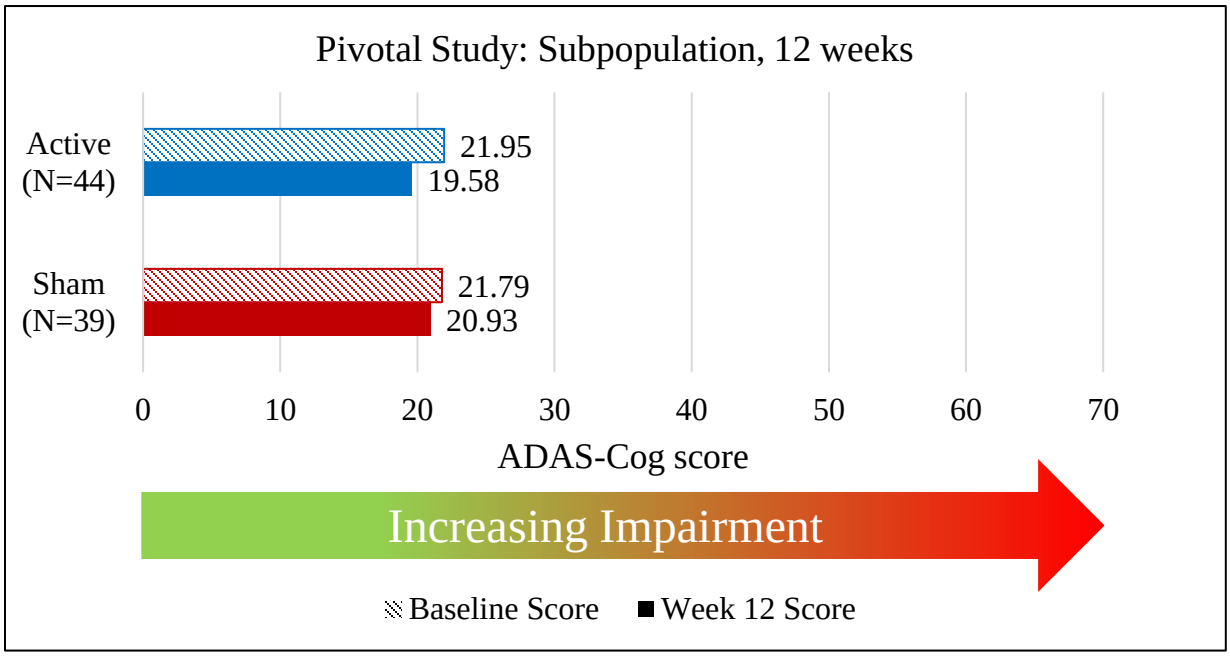


Figure 32. Pivotal Study Post-Hoc Analysis Results, Subgroup, ADAS-Cog Groups Means at Baseline and 12 Weeks

In this subgroup, the CGI-C at the 7-week timepoint resulted in a difference of 0.07 in favor of the treatment group. At 12 weeks, this difference reaches 0.40 in favor of the treatment group. Here again at this 12-week timepoint, the 0.40 difference between groups is the highest magnitude CGI-C difference in favor of the treatment group.

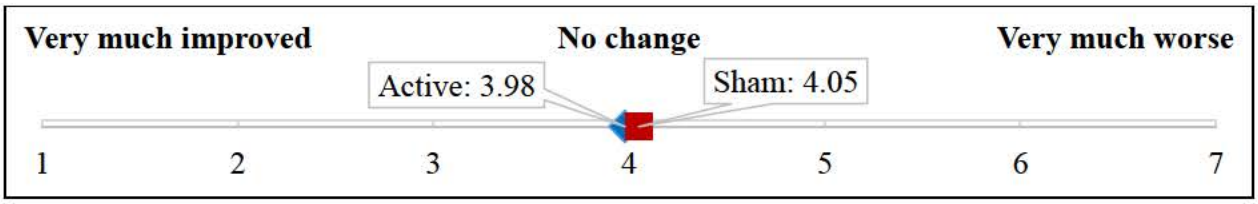


Figure 33. Pivotal Study Post-Hoc Analysis, Subgroup, CGI-C at 7wks

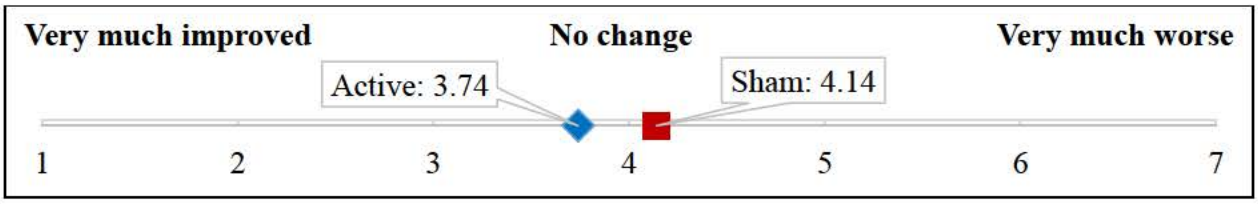


Figure 34. Pivotal Study Post-Hoc Analysis, Subgroup, CGI-C at 12wks

FDA Summary Comments on Pivotal Study Post-Hoc Analysis ADAS-Cog≤30 Subgroup

FDA works with sponsors to design prospective, well controlled studies to evaluate the safety and effectiveness of medical devices. Prospective, controlled studies are an effective method to independently and objectively assess the safety and effectiveness of medical devices. Pre-specified hypotheses that are documented and planned to be tested before any examination of the data are a tenet of good trial design. In the sponsor's case, the analysis of an impact on disease severity at baseline was pre-specified as a covariate, but the hypothesis that it would represent an independent cohort that demonstrates a larger and more consistent treatment effect was not. That is, despite the interaction test being pre-specified, the sponsor's intention was to make a claim for the effectiveness of neuroAD over sham for the entire population. The SAP did not further specify methods of analysis if the test was statistically significant (but, the overall averaged effect was not). Therefore, changing the intended population after analyzing the data and finding that the overall test was not significant, amounts to a post-hoc hypothesis test. Therefore, analyses associated with the subgroup defined by baseline ADAS-Cog <= 30 carry much greater uncertainty. These types of post-hoc analyses are generally considered to be exploratory and hypothesis-generating.

In contrast to prospective hypothesis tests, a data-driven hypothesis is generated and tested after an examination of the data. After finding the primary endpoint to be non-significant, an investigator might present a nominally significant post-hoc analysis of a subgroup as a substitute for the primary endpoint analysis. However, the usual calculation of type I error may be incorrect, especially when the data themselves suggest the hypothesis test. The same data should not be used both to generate a new hypothesis and to test it (Piantadosi, 1997). In this case, the hypothesis is that patients with baseline ADAS-Cog <=30 benefit from the neuroAD device over sham. Since the pivotal study results were used to generate that hypothesis, an independent dataset should generally be used to provide verification.

Post hoc analyses expose the risk of approving medical devices that have no beneficial effect. An apparent treatment difference discovered after an unplanned post-hoc analysis may be due to coincidence. This phenomenon can manifest itself in post-hoc subgroup analyses, done in the hope of discovering differences that support one treatment over another. Post-hoc analyses are also problematic because they neither apply to the intention to treat population in the study nor conform to the randomization model of statistical inference (especially if randomization was not stratified by the subgroup).

For these reasons, FDA typically depends on well-controlled studies with prospectively defined statistical analysis plans to best determine the safety and effectiveness profile for any given device. In this instance, the sponsor has provided an exploratory post-hoc analysis of the Neuronix pivotal study and pooled analyses from supplemental investigations. The sponsor also provided data it believes is confirmatory from two independent studies in Korea (Korean Pilot and interim data from Korean Pivotal). The panel will be asked if this evidence is sufficient to confirm the subgroup identified in the pivotal investigation and demonstrate device safety and effectiveness.

FDA Analysis of Sponsor's Post-Hoc Subgroup

As presented in the section above, Neuronix explored a trend in the pivotal study data that demonstrated a higher difference in favor of the treatment in a subset of the patients who had a baseline ADAS-Cog≤30 at a secondary endpoint assessment timepoint of 12 weeks. It is intended that the subgroup of subjects in the pivotal study that had a baseline ADAS-Cog score that was ≤ 30 define a clinically plausible subset of mild to moderate Alzheimer's disease patients for which the device is superior to sham. This subset eliminated 8 subjects from the treatment group and 8 subjects from the sham group with baseline ADAS-Cog scores > 30.

Please refer to the Sponsor Executive Summary for information about how this subgroup was chosen as this subgroup was derived independently from FDA input. The following are analyses performed by FDA to independently investigate this proposed subgroup.

After finding a post-hoc subgroup that shows a stronger observed treatment effect than the overall set of subjects, it is recommended to check balance across the treatment groups on other baseline covariates to make sure that the observed treatment difference is not due to an imbalance on another covariate that is also related to response. Using data from the sponsor, the FDA looked at summary statistics across treatment groups within the subgroup of ADAS-Cog baseline <= 30. Table 14 shows baseline means for available covariates. There do not appear to be imbalances that could be responsible for an enhanced observed treatment effect within the subgroup. In addition, interaction tests using these other baseline covariates did not show interactions with treatment.

Table 14. Baseline Means for Subjects with Baseline ADAS-Cog<=30

	Active (n = 45)	Sham (n = 40)
ADAS-Cog	21.91	21.90
MMSE	22.27	21.60
Age (years)	78.24	75.7
Gender – Proportion Male	25/45 = 0.56	24/40 = 0.60
Proportion on no AD Medication	12/45 = 0.27	8/40 = 0.20

A test of the interaction between ADAS-Cog baseline and treatment group on change in ADAS-Cog at the primary time point was pre-specified in the statistical analysis plan (SAP) to be conducted in the primary analysis at 7 weeks. Neuronix reported that the test was statistically significant, implying that a potential treatment effect differed depending on the value of ADAS-Cog at baseline. Here, we also discuss the interaction at 12 weeks. For both 7 and 12 weeks, the interaction has a similar interpretation: that is, that the Active group outperforms the Sham on average at lower baseline ADAS-Cog values (milder AD), but as baseline ADAS-Cog increases (worse AD), Sham outperforms Active.

Figure 35 and Figure 36 show graphical interpretations of the detected interactions at 7 and 12 weeks. Circles represent actual observations. Solid lines show the fitted regression lines from the model where the interaction terms were used[3]. For the 7-week time point, the model says that Active outperforms Sham on average at baseline ADAS-Cog values less than 20, with Sham out-performing Active after 20. For the 12-week time point, the model says that Active outperforms Sham on average at baseline ADAS-Cog values of either less than 24 (for linear model) or less than 29 (for smoothed-fit model).

[3] Dotted lines show smoothing spline fits through the observations. The dotted lines are meant to detect whether a linear model is adequate to explain the interactions. Although there appears to be some evidence of nonlinearity, the relative patterns of the dotted lines by color is similar to the patterns of the solid lines by color, so that a linear model is appropriate.

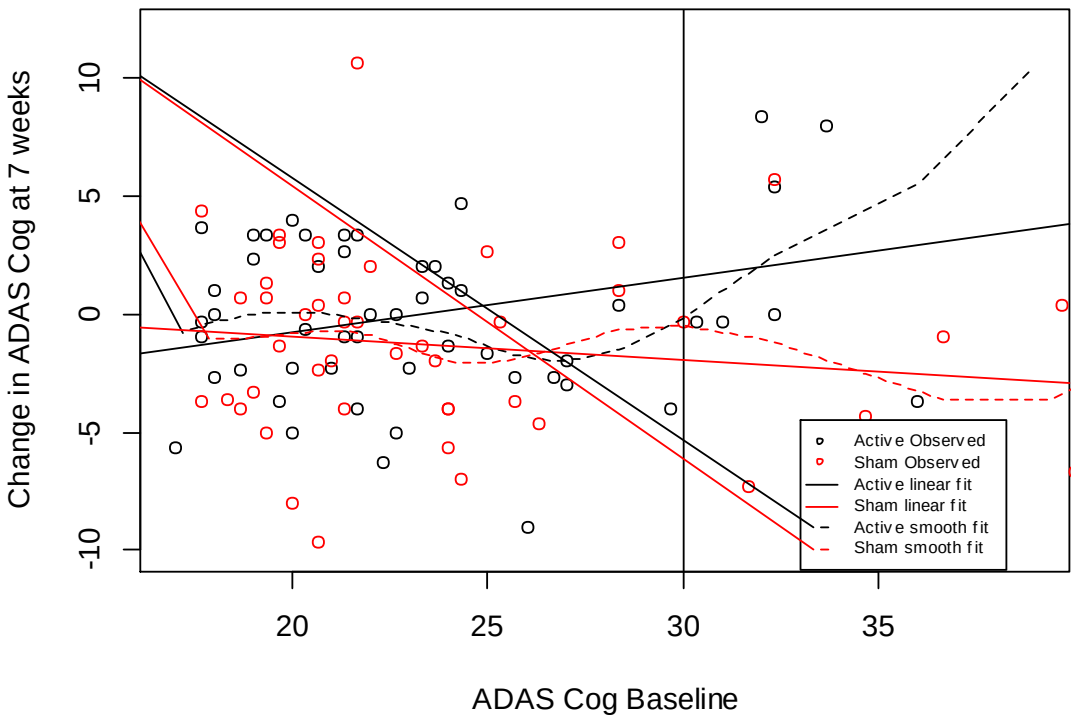


Figure 35. Pivotal Study - Detected Interaction at 7 Weeks

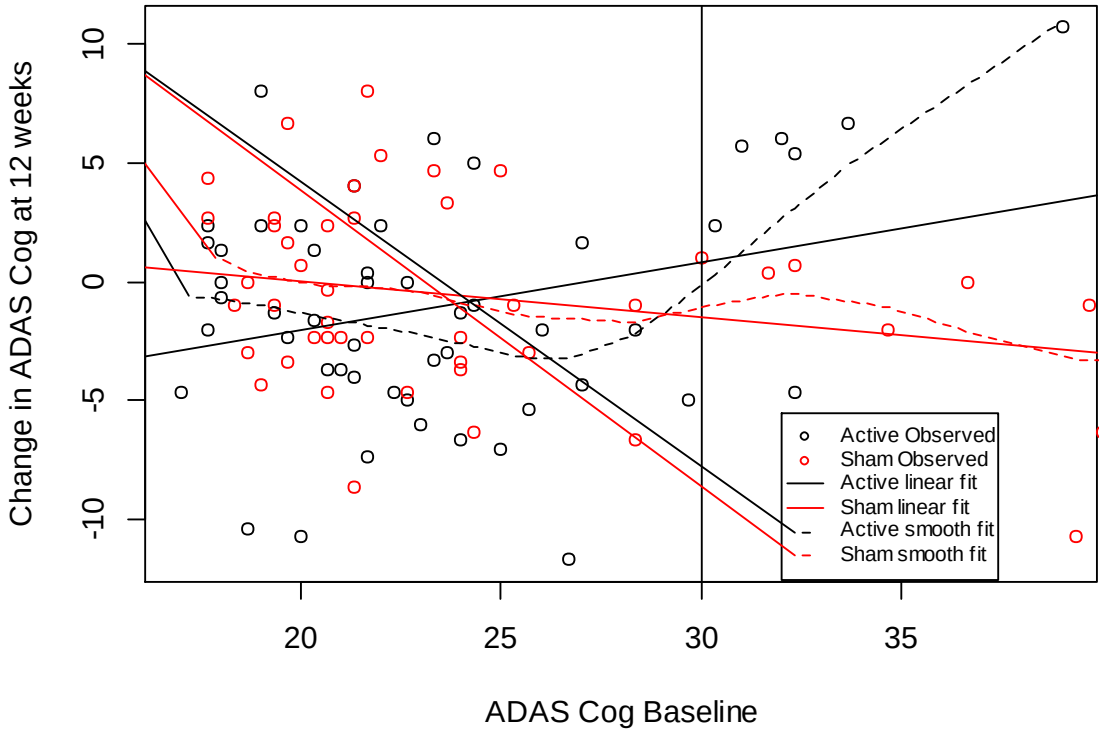


Figure 36. Pivotal Study - Detected Interactions at 12 Weeks

Table 15 shows the observed mean differences for subjects who met a baseline ADAS-Cog threshold of up to 18 to 45, by assessment time point. The sample sizes are cumulative. At both time points, the mean difference increases (favoring sham more) as subjects with higher baseline values are included. At 12 weeks, the mean difference remains in favor of Active, despite increasing over baseline thresholds.

Table 15. Observed Mean Difference on ADAS-Cog by Baseline ADAS-Cog Score, 7 and 12 week timepoints

7-week endpoint							
Baseline ADAS-Cog	<=18	<=20	<=25	<=30	<=35	<=40	<=45
Mean difference (Active-Sham) (95% CI)	-1.05 (-39.0, 37.0)	0.85 (-1.90, 3.59)	1.01 (-0.66, 2.68)	0.47 (-1.05, 1.99)	1.02 (-0.53, 2.56)	1.63 (-0.06, 3.32)	1.45 (-0.27, 3.17)
n (active/sham)	7/2	15/13	38/34	45/40	51/43	53/47	53/48
12-week endpoint							
Baseline ADAS-Cog	<=18	<=20	<=25	<=30	<=35	<=40	<=45
Mean difference (Active-Sham) (95% CI)	-3.78 (-7.26, -0.31)	-1.64 (-4.96, 1.68)	-1.46 (-3.40, 0.47)	-1.61 (-3.39, 0.17)	-0.95 (-2.68, 0.78)	-0.35 (-2.13, 1.43)	-0.42 (-2.19, 1.35)
n (active/sham)	7/2	14/13	37/34	44/39	50/42	51/46	51/47

Despite the observed evidence in the US pivotal study that a milder baseline ADAS-Cog benefits from the device, and that a less mild ADAS-Cog will not, it is not clear at what baseline score the benefit would be expected. Neuronix proposed a baseline score of 30 on the ADAS-Cog as cut-off for benefit using the neuroAD. Based on the univariate analyses above, this value appears reasonable. However, as stated above, for exploratory subgroups that are found to be hypothesis-generating, we recommended that the subgroup be validated or "confirmed" on an independent data set. Otherwise, the risk of type I error inflation or exaggeration of treatment effect is too high because the data dictate the cut-off.

The sponsor has discovered a qualitative interaction at 12 weeks, after an overall averaged effect has been found to be of low magnitude (-0.42). They proposed to divide the data set into two groups based on a baseline ADAS-Cog value of 30. With values <= 30 (N=85), the mean difference between groups is -1.61 (-3.35, 0.13), favoring Active. With values > 30 (N=16), the mean difference between groups is 6.61 (1.96, 11.26), favoring Sham. The US pivotal study shows general inconsistency in averaged treatment effect compared to the supplemental studies (see Appendix V. Post Hoc Analyses Using Supplemental Clinical Data Sources). In addition, each of the supplemental studies is small, resulting in significant uncertainty in the interpretation of the results.

Therefore, it is important that the sponsor validate (in another data set) that superiority of neuroAD over sham in patients with baseline ADAS-Cog <= 30 because there is no apparent treatment effect in the whole cohort of the sponsor's US pivotal study. While the sponsor presents interim data from the Korean pivotal study in a very small cohort of patients (22 treated patients, 11 active) with ADAS-Cog<30, this is not sufficient evidence to serve as a confirmatory group.

FDA Validation of Post-Hoc Subgroup

FDA performed an analysis to find a subgroup that is predictive of responding more favorably to Active over Sham in order to investigate the sponsor's proposal of a post-hoc subgroup. We attempted to validate both the ADAS-Cog baseline <=30 cut-point as a predictive marker for device success, as well as a more comprehensively chosen subgroup using several baseline covariates included in a data set provided by the sponsor. Because of the presence of several external data sets (please see Appendix V. Post Hoc Analyses Using Supplemental Clinical Data Sources) that used the sponsor's device compared to sham, FDA was also able to conduct a rudimentary external validation.

Although an external independent dataset is preferred for validation, there may be reasons for doing an internal validation if the external data sets are unavailable, not representative of the patient population, or not similar to the pivotal study. A key argument of the sponsor is that the studies (at least the Korea studies) are comparable to the pivotal study, and that their results represent the device's effectiveness. Indeed, their stated reason for providing a meta-analysis of the set of studies in their executive summary is the similarity of study designs. However, in case that argument is not correct, an internal validation can provide an adjusted estimate of the treatment effect, taking into account that a post-hoc finding may be a random high as a result of looking for an observed treatment effect. Nonetheless, an internal validation essentially uses data from the same study that was used to "discover" the subgroup post-hoc, even if it doesn't use exactly the same patients. Therefore, its usefulness is limited.

Internal Validation of Post-Hoc Subgroup

An internal validation typically resamples observations from the study data in order to assess a post-hoc subgroup. The resampling can either be with replacement (e.g., bootstrapping) or without replacement (e.g., cross-validation). In either case, many repeated resamples are taken and then averaged to obtain the adjusted estimate. Associated standard errors can be obtained using additional methods. From each resampled data set, a subgroup is determined such that the treatment effect is larger in the subgroup than in the overall study cohort. The subgroup is then "validated" on study data that were not used in that resample. In this sense, the validated data are not the exact values used to determine the subgroup. Appendix VII contains the statistical details of the internal validation that was done by FDA.

The baseline covariates used to discover a subgroup in the internal validations were either 1) baseline ADAS-Cog alone or 2) baseline ADAS-Cog, baseline MMSE, gender, whether AD medication was reported as taken at the baseline visit, and age. Table 14 shows the result of 1) and 2) using a bootstrapping method to obtain an adjusted estimate of the subgroup-specific treatment effect. The "adjustment" is intended to correct for potential spurious values as a result of bias from a post-hoc choice. In general, for 1), the average cut-off value for baseline ADAS-Cog was 30.07; for 2), cut-offs were most often chosen for baseline ADAS-Cog (usually around a value of 30), baseline MMSE (around a value of 22.5), and gender.

Table 16 shows that if baseline covariates are assumed to be truly predictive of the effectiveness of neuroAD over sham, an adjusted estimate of its effect is around -1.53, considering *all* baseline covariates for potential splits. If we only consider baseline ADAS-Cog, then an adjusted estimate of neuroAD's effect over sham is -1.33. This estimate is lower than the data-based estimate of -1.61 obtained by the sponsor. More importantly, in Table 16 we also include an estimate of the excess treatment effect in the subgroup over the entire population. The data-based estimate from the sponsor is -1.19 (i.e., -1.61 – (-0.42)). This estimate is roughly similar to the internally validated estimates shown in Table 16, which range from -0.76 to -1.21.

Table 16. Internal Validation of "Predictive" Subgroup

	Bootstrapped Bias-corrected estimate of treatment effect (using cut-offs on all baseline covariates)	Bootstrapped Bias-corrected estimate of treatment effect (using cut-offs on baseline ADAS-Cog)
In predictive subgroup	-1.53 (-2.61, -0.45)	-1.33 (-2.33, -0.43)
Enhanced difference of predictive subgroup over the entire population	-1.21 (-1.92, -0.50)	-0.76 (-1.46, -0.06)

External Validation of Post-Hoc Subgroup

Whereas internal validation can be useful when an external dataset is not available, there are still inherent biases in any adjusted estimate from an internal validation. The internal validation is conducted under the assumption that the study was run free of any study-related or investigator-related biases. *If any biases are present in the pivotal study results, they will also be present in the cross-validation or bootstrapped adjusted result*. Therefore, it is often better to validate potentially predictive subgroups by using an independent data set. Fortunately, the sponsor provided several small studies that could be used as one or more validation data sets. However, all of the studies have small sample sizes, and even fewer subjects are present in the proposed subgroup. According to the sponsor, the studies that are appropriate for assessing external validation of the post-hoc subgroup are the Korea pilot study and its interim pivotal version (Korea-2). The Korea-2 only contains subjects who meet the baseline ADAS-Cog <= 30 cut off, as it enrolled mild AD patients as determined by baseline MMSE. A small study conducted in Italy was an RCT that used both NeuroAD and a similar sham as did the US pivotal study. Therefore, due to the small size of the Korea pilot study, we also combined it with the Italy study in order to increase the number of subjects.

Table 17 contains results from this external validation of the baseline ADAS-Cog<=30 subgroup. The three estimated mean differences in the first row of the table show an average treatment effect in the baseline ADAS-Cog <= 30 subgroup that favors neuroAD over sham. However, it is not clear that the subgroup shows an enhanced effect over the entire study. It does in the Italy study, but not in the Korea pilot. Combining the two studies does not show an enhanced effect, as it is dominated by the Korea pilot study. The "enhanced differences" are subject to variability, which was not estimated, and therefore the values should not be interpreted as certain. Because of this uncertainty, it is not clear that the sponsor's proposed restriction will generalize to the population at large.

Table 17. External Validation of baseline ADAS-Cog<=30

External Validation (of baseline ADAS-Cog <= 30)			
	Korea Pilot Study (14 Active; 7 Sham)	Italy (5 Active; 1 Sham)	Both RCTs combined (Korea+ Italy) (19 Active; 8 Sham)
In subgroup	–1.79 (–5.39, 1.81)	–1.80 (NA)	–1.24 (–4.50, 2.02)
In entire study	-2.51	+0.17	-2.03
Observed enhanced difference of subgroup over the entire population	–1.79 – (–2.51) = +0.72	-1.80 – (+0.17) = -1.97	-1.24– (-2.03) = +0.79

Note that Table 17 excludes the Korea Pivotal (Korea-2) study, which only enrolled patients with baseline ADAS-Cog <= 30. Its subgroup estimate is -1.73 (-4.74, 1.28). This is also the estimate in the entire study; so, the observed enhanced difference would be identically 0.

Hierarchical Model adjustment of subgroup-specific treatment effect (Table 18) and comments on Sponsor's meta-analysis (Table 19)

Another way to get an adjusted estimate of a subgroup-specific treatment effect is to fit a hierarchical model where individual subgroup estimates are revised via shrinkage estimation. Shrinkage estimation imposes a multiplicity adjustment so that potentially spurious extreme estimates are shrunk toward more reasonable values. With post-hoc subgroup formation, shrinkage estimation could provide a more "honest estimate" of the actual treatment effect in the chosen subgroup, had it not been selected based on results (provided that subgroup separation is clinically meaningful). The table below provides the hierarchical model estimates of the treatment differences within each subgroup, as well as the sponsor's separate, non-hierarchical estimates. The hierarchical model estimate for the subgroup-specific treatment differences are reduced in magnitude. The estimated posterior probability that Active is better than Sham in the baseline <= 30 subgroup is 0.85.

Table 18. Non-Hierarchical versus Hierarchical Estimates of Treatment Differences within Each Subgroup

	Non-hierarchical subgroup-specific estimates (SD) (95% CI)	Hierarchical model subgroup-specific estimates (SD) (95% CI)
Treatment difference (ADAS Cog baseline<=30)	-1.61 (0.89) (-3.35, 0.13)	-1.39 (0.95) (-3.26, 0.46)
Treatment difference (ADAS Cog baseline > 30)	6.61 (2.37) (1.96, 11.26)	5.34 (2.47) (0.05, 9.93)

In their Executive Summary, the sponsor performed a meta-analysis that included their US pivotal study, along with the pilot study done in Korea, as well as another study (called a "pivotal" study) done in Korea. Because the protocols were similar among the three studies, the sponsor assumed that the studies could be considered exchangeable, and therefore included together within a meta-analysis. The assumption of exchangeability may be questioned due to subjects being enrolled in different countries with different cultures across the three studies. Also, according to the sponsor there were differences in baseline motor thresholds (MT) across the studies. In the Korean pilot study, for example, the average MT at baseline for all active patients was 94.4%, which is significantly higher than the average MT at baseline for all active patients from the neuroAD pivotal study (74.2%). Also, for the 4 active patients in the Korean pilot study who had baseline ADAS-Cog > 30, the average motor threshold was 93%. Because the intensity of TMS is typically adjusted according to the patient's baseline MT, the intensity of TMS treatment may differ between the Korea pilot study and the US pivotal study.

In addition, the sponsor's meta-analysis excluded any subjects in the three studies whose baseline ADAS-Cog value was > 30. The weighted mean estimate that the sponsor obtained for the US study is reproduced in Table 19 below (row 1). The observed result from the US study is also provided. However, excluding the poorer performing subgroup (i.e., baseline ADAS-Cog > 30) essentially becomes a separate analysis of a subgroup, treating the subgroup samples as though they were the only subjects enrolled in the respective studies. Instead, a subgroup analysis should contain two types of variation: variation in the true subgroup-specific treatment effects and variation from random sampling of the patient outcomes. Because subgroups are smaller than the entire study, their sampling variation can be high. Also, their results may be subject to selection bias if the results are noticed only because they are

favorable. Due to such influences, post-hoc subgroup estimates may be exaggerated over their true values. There exist many ensemble methods that can provide "shrinkage" estimates of subgroup-specific treatment effects so that the estimates better reflect the actual treatment effects by reducing the random sampling variability. These methods require that data from the "unpromising" subgroup also be included in the analysis. (Pennello & Rothmann, 2018)

Table 19 contains FDA's adjusted subgroup-specific treatment effect estimates in the US study, using a Bayesian hierarchical method from section 10.7 in Pennello and Rothman (2018). Compared to the observed result (row 2), which was obtained by calculating the group mean difference within each subgroup separately, the adjusted estimates are shrunk toward less extreme values. The treatment difference in the baseline ADAS-Cog <= 30 subgroup for the US study is shrunk to -1.26 (-2.93, 0.45), and the treatment difference in the baseline ADAS-Cog > 30 subgroup is shrunk to 2.15 (-1.95, 6.17). The estimated posterior probability that the ADAS-Cog <= 30 subgroup difference is less than 0 is 0.93.

Table 19. Sponsor Estimates versus FDA Estimates for the US Pivotal Study (including US study, Korea pilot, and Korea-2 studies within a hierarchical model)

	US Estimate in baseline ADAS-Cog <=30 subgroup	**US Estimate in baseline ADAS-Cog > 30 subgroup**
Sponsor's Pivotal Study Subgroup Meta-Analysis	-1.61 (-3.36, 0.14)	NA
Observed result	-1.61 (-3.35, 0.13)	6.61 (1.96, 11.26)
FDA's hierarchical model analysis	-1.26 (-2.93, 0.45)	2.15 (-1.95, 6.17)

Conclusion of Post-Hoc Analysis ADAS-Cog≤30 Subgroup

Our analyses suggest that if baseline ADAS-Cog <= 30 is clinically meaningful as predictive of treatment benefit by neuroAD over Sham, its average effect is about 1.26 – 1.40 points better on the ADAS-Cog at 12 weeks, representing about a 1.0 point average enhanced improvement over the entire population (second row of Table 16). However, this subgroup could not be successfully confirmed in independent studies provided by the sponsor.

Concerns regarding the statistical methods result in significant uncertainty in the results of the post-hoc analysis. These factors notwithstanding, we looked at the results of the sponsor's post-hoc analysis limited to the subgroup of patients with a baseline ADAS-Cog score ≤ 30 in terms of clinical meaningfulness. Here again at the 7 week timepoint the results favor the sham group by a +0.47 point difference between groups. At the 12 week timepoint this difference between groups favors the treatment group by -1.61 points. Although this is a larger magnitude change than the overall cohort at the same timepoint, it still does not appear to reach a minimum clinically meaningful difference on the ADAS-Cog scale based on the review team's assessment. This assessment is discussed further in the Benefit-Risk Assessment section.

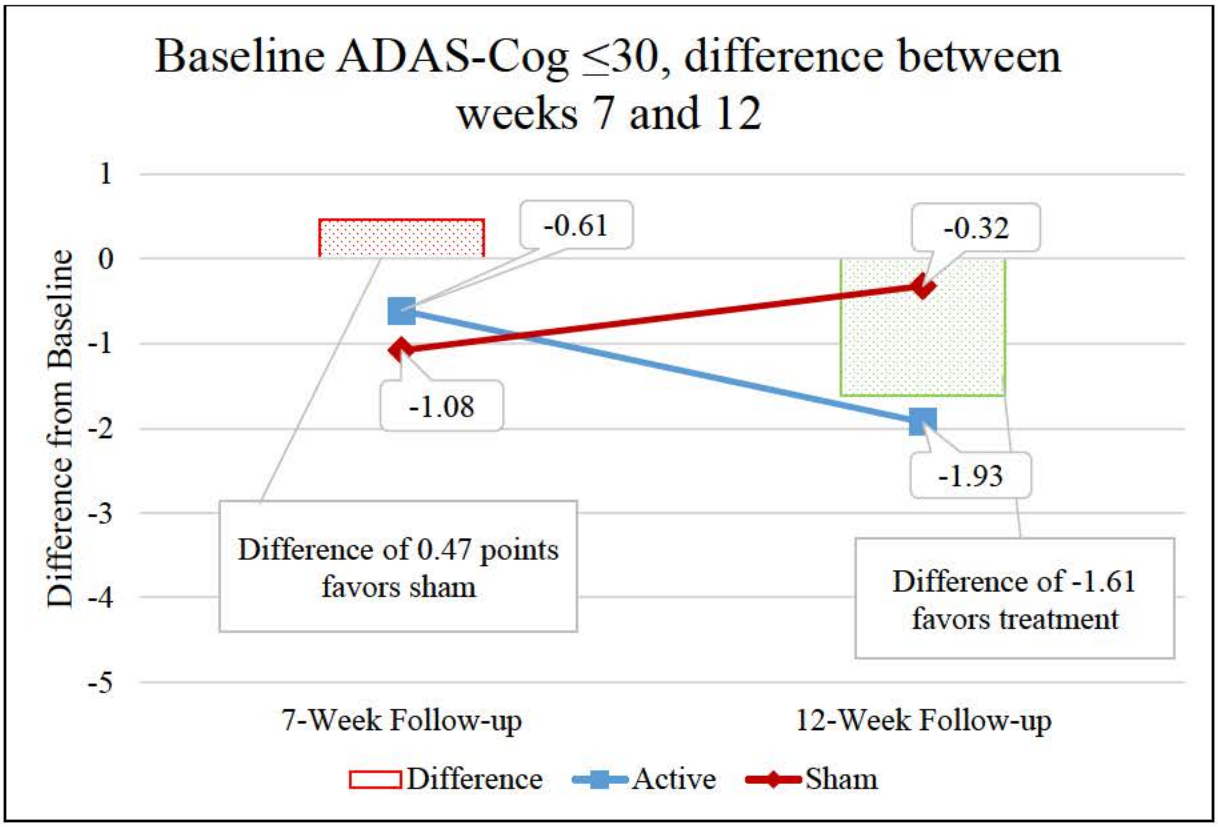


Figure 37. Pivotal Study Post-Hoc Analysis – Mean Change in ADAS-Cog Summary Results

As with the entire cohort, FDA looked to the secondary assessment scale, the CGI-C, to provide clinical context to the ADAS-Cog results. At the 7-week assessment timepoint, the difference between groups in this cohort of baseline ADAS-Cog ≤30 was 0.07. At 12 weeks the change increased to 0.4 in favor of treatment (the highest magnitude of change shown in the CGI-C between groups was found in this post-hoc subgroup at 12 weeks). These small changes in CGI-C scale around 4 (no improvement) provide little if any clear clinical support to the ADAS-Cog results.

In conclusion, the analysis of an interaction between baseline ADAS-Cog and treatment group on 7-week change from baseline was pre-specified. However, the hypothesis of a particular cut-point to use on baseline ADAS-Cog to restrict the indication was not proposed before the study was finished. Baseline ADAS-Cog may be relevant with respect to a patient achieving superior benefit from neuroAD over sham. However, further input is needed from the panel on whether the cut-point of 30 on this assessment has been shown to be an appropriate predictor of effectiveness, given the following observations:

1. The cut-point was chosen by examining the study data itself, and then tested on those data.
2. FDA was not able to verify that the supplemental studies showed an enhanced benefit in patients meeting the cut-point over the whole study cohort. The supplemental studies collectively showed different patterns of treatment benefit over sham than did the US study across the two follow-up assessments, as well as across the ADAS-Cog<30 and ADAS-Cog>30 subgroups. The supplemental studies had small sample sizes, but even pooled together they showed a different pattern that presents uncertainty in concluding device effectiveness in a US population. The supplemental studies collectively showed different patterns of treatment benefit over sham than did the US study at 7 weeks follow up.

<u>Panel Question</u>: The Panel will be asked to discuss and make recommendations on whether the ADAS-Cog≤30 population is a clinically plausible subset.
<u>Panel Question</u>: The Panel will be asked to discuss and make recommendations on whether the post-hoc identification of the ADAS-Cog≤30 population at a later time point when no treatment is given is an adequate analysis of the pivotal study data, in concert with the supplemental data provided, to demonstrate probable benefit.

Additional Analyses using Pivotal Study Data

ADAS-Cog Waterfall Plot

To better understand the proposed ADAS-Cog≤30 subgroup, the FDA requested the following waterfall graphs from Neuronix. The waterfall graphs below show the entire cohort of the pivotal study and provide information regarding the baseline ADAS-Cog score and the final ADAS-Cog score at 7 weeks and 12 weeks. As denoted by the white-filled bars, the majority of the cohort of subjects with an ADAS-Cog score > 30 appear to be poor-performing active subjects (positive changes on the ADAS-Cog scale) and high-performing sham subjects (negative changes on the ADAS-Cog scale). Based on this view of the data, it appears that limiting the data to only those with a baseline ADAS-Cog score ≤ 30 would influence the results in favor of the treatment from both the active and sham eliminations.

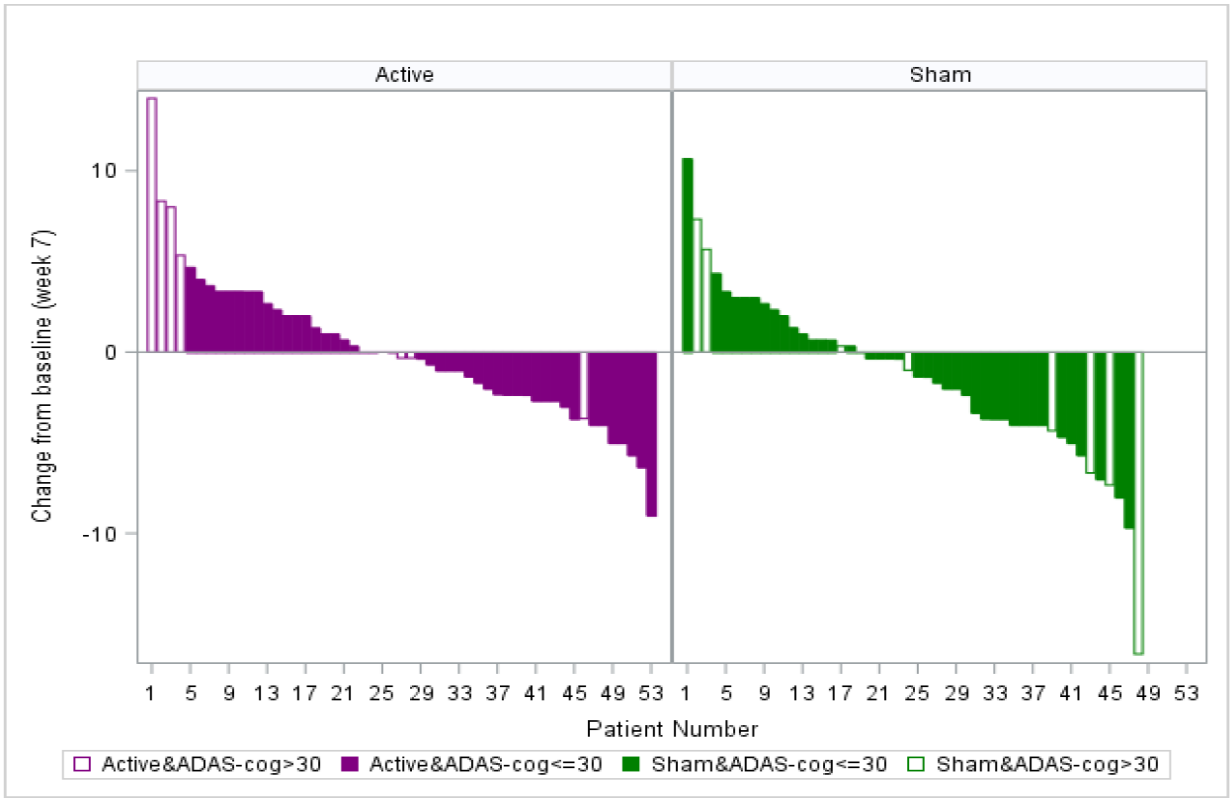


Figure 38. Pivotal Study Post-Hoc Analysis - Neuronix Waterfall Plot, 7 wks

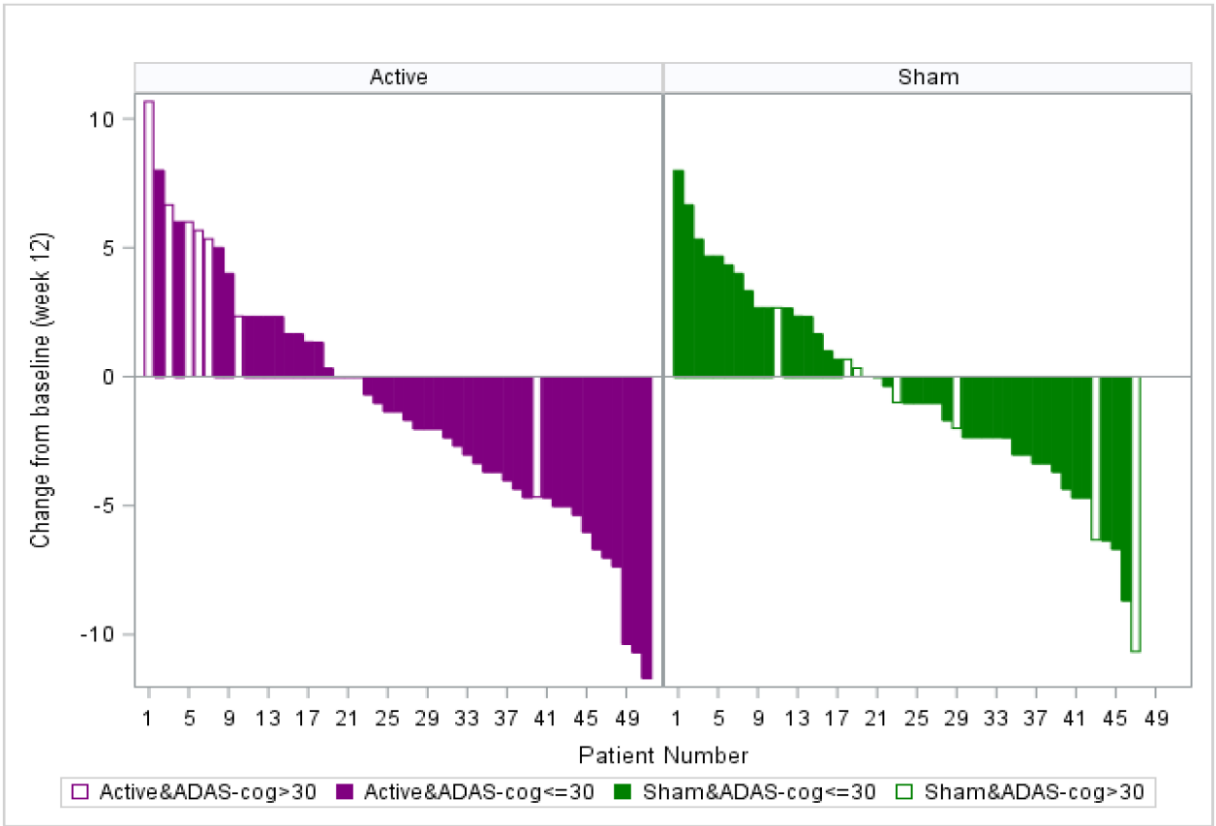


Figure 39. Pivotal Study Post-Hoc Analysis – Neuronix Waterfall Plot 12 wks

Pivotal Study ADAS-Cog Responder Analysis

Neuronix has presented post-hoc responder analyses for the ADAS-Cog and the CGI-C results. When reviewing these analyses, we caution the panel that there was no agreed-upon responder rate that was pre-specified for either of these scales. Furthermore, we are unaware of a standard definition of a responder on the ADAS-Cog or the CGI-C assessment scales.

While we present our own responder analysis for the ADAS-Cog results in this section as a companion to those presented by the sponsor, we do not present the CGI-C. Instead of viewing the results of the CGI-C assessment in terms of "responders" we recommend that the panel refer to the individual breakdown of the CGI-C responses as the most informative way to categorize the responses on the CGI-C scale (e.g., Figure 22). We recommend this method over the responder analysis because it is unclear how to categorize a responder on the CGI-C scale for this trial design. For example, the sponsor has included subjects with scores from 1-4 as responders. It is unclear how a score of "4" (no change) on the CGI-C can be considered a responder to the neuroAD intervention when this subject was determined to have no change from a clinical perspective. As the neuroAD is intended as a treatment intervention, for mild to moderate AD, the CGI-C assessments were conducted serially over a period of 12 weeks (baseline, week 7 and week 12) within approximately six weeks of each other. The clinical trial was not designed to assess for the modification of disease progression and the twelve-week time interval of the study is insufficient to comment on prevention of disease progression. In summary. due to the natural course of AD and the clinical trial design, a score of no change at 7 weeks or 12 weeks may not be indicative of device performance.

While the concerns regarding viewing the results of the ADAS-Cog in the form of "responders" remain, we provide S-curve plots of the entire pivotal study cohort and the baseline ADAS-Cog<=30 subgroup at 7 and 12 weeks (using PE population) to compare with those the sponsor has provided in their Executive Summary. The information below presents the percentage of patients achieving a certain degree of change on the ADAS-Cog along a continuum from baseline to 7 weeks and baseline to 12 weeks. We choose to highlight a "responder rate" of -3-points as this was previously proposed as a minimum clinically important difference (MCID) on the ADAS-Cog scale to the sponsor in the AINN Letter.

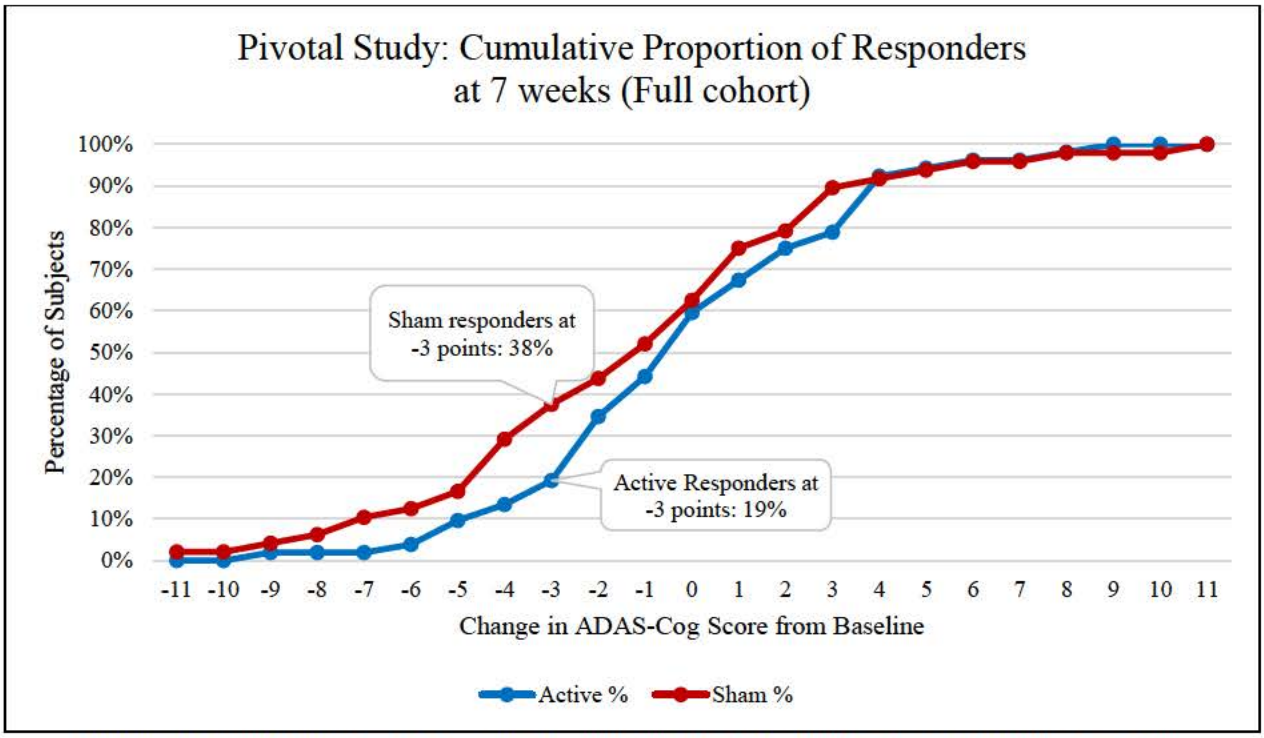


Figure 40. Pivotal Study ADAS-Cog Responder Analysis, Full Cohort, Baseline to 7 Wks

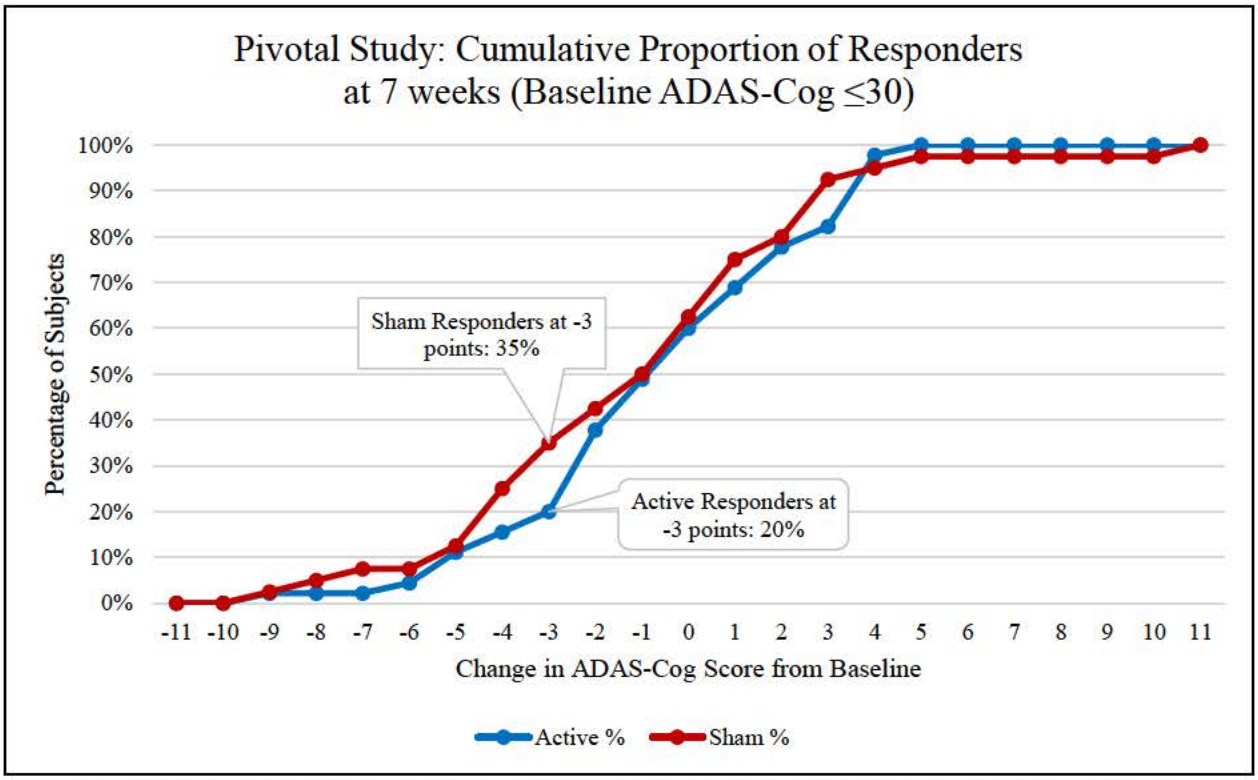


Figure 41. Pivotal Study ADAS-Cog Responder Analysis, Subgroup, Baseline to 7 Weeks

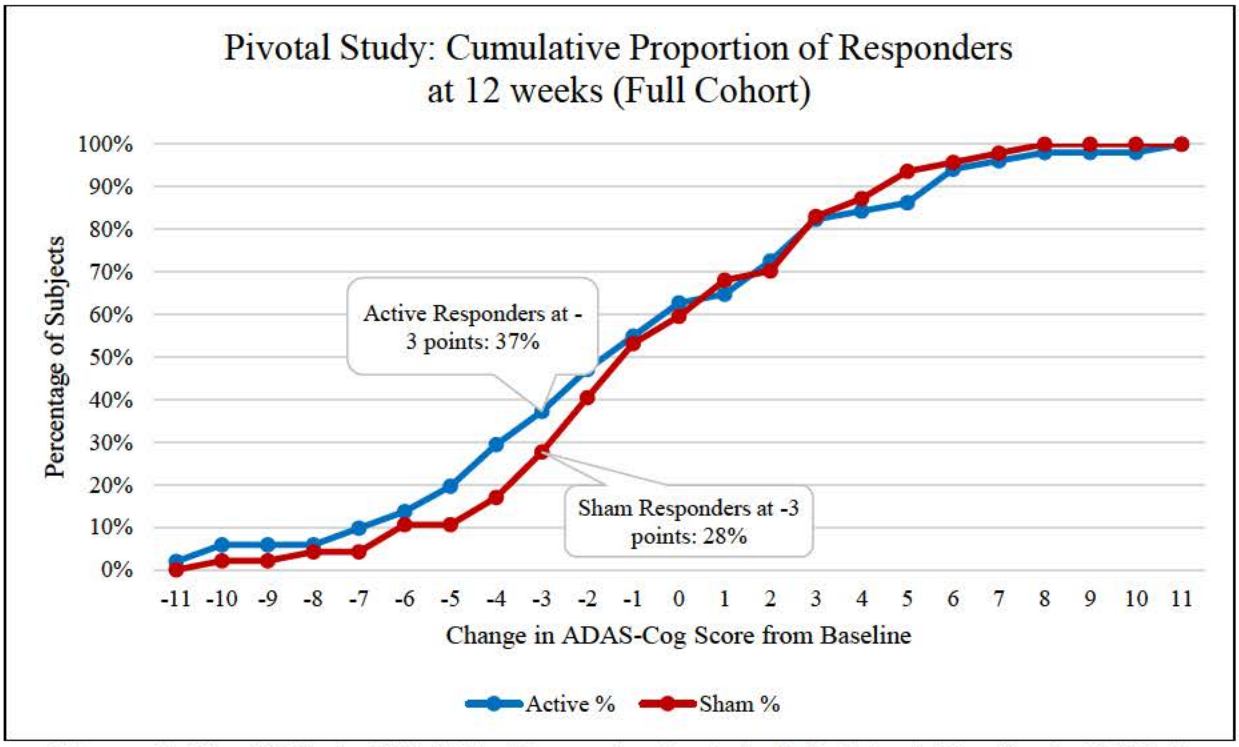


Figure 42. Pivotal Study ADAS-Cog Responder Analysis, Full Cohort, Baseline to 12 Weeks

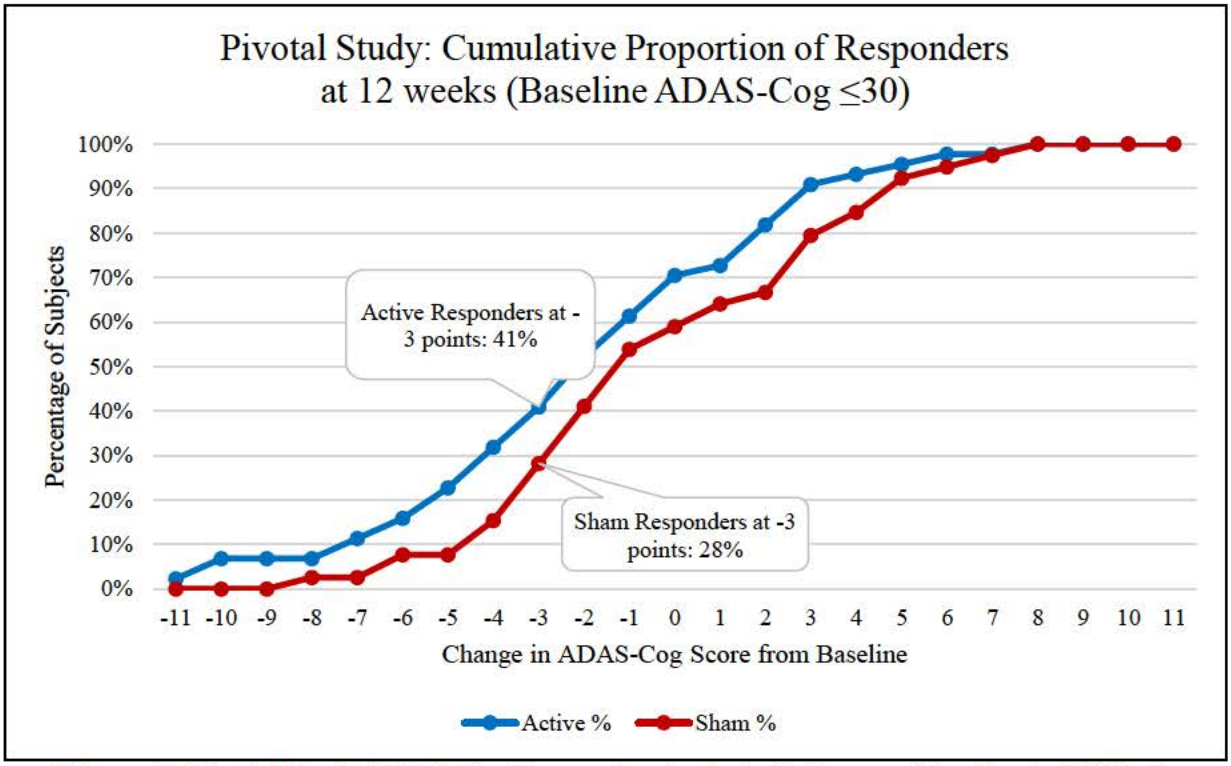


Figure 43. Pivotal Study ADAS-Cog Responder Analysis, Subgroup, Baseline to 12 Weeks

Benefit-Risk Assessment

Panel Question: The Panel will be asked to discuss and make recommendations regarding whether the probable benefits to health outweigh the probable risks.

FDA is committed to providing timely patient access to novel devices that are demonstrated to provide a reasonable assurance of safety and effectiveness. In assessing benefits and risks, CDRH weighs several factors as outlined in the FDA guidance document, "Factors to Consider When Making Benefit-Risk Determinations in Medical Device Premarket Approval and De Novo Classification" ("Benefit-Risk Guidance"). These factors include, but are not limited to the type of benefit(s), the magnitude of the benefit(s), the likelihood of patients experiencing one or more benefits, duration of effects, patient perspective on benefit (when known), the severity of harm, likelihood of risk, and uncertainty regarding the safety and effectiveness.

FDA is seeking the panel's input on whether the available evidence demonstrates a favorable benefit-risk profile for the neuroAD device for the proposed indication. In particular, the Agency acknowledges that the risks of the device are low. However, the pivotal study failed to meet its prespecified primary endpoint and reliance on post hoc analyses introduces significant uncertainty in the results. Therefore, FDA will be seeking input on several issues, including whether: 1) the available data demonstrate device effectiveness; 2) there is evidence of a clinically meaningful benefit; and 3) the overall benefit-risk assessment of the neuroAD for the proposed indication is favorable.

The following sections provide more information regarding FDA's assessment of the benefit-risk ratio of this device.

FDA Summary of neuroAD Benefit

Device Effectiveness

The type of benefit that is purported for the neuroAD is a relief of cognitive symptoms of Alzheimer's dementia. The primary endpoint, ADAS-Cog, was chosen to focus specifically on the cognitive symptoms of this disease. The secondary endpoint, CGI-C, was meant to compliment and capture global clinical changes to support the cognitive endpoint but was not intended for labeling claims. There were no other disease symptoms targeted. The subgroup of baseline ADAS-Cog≤30 was not pre-specified, but was identified through post-hoc analysis.

As the largest and most rigorously-designed study of the device to date, FDA believes that the pivotal study primary endpoint (change in ADAS-Cog at 7 weeks, regardless of baseline ADAS-Cog measurement) should carry the most weight. As noted above, this endpoint favors the sham group by +1.45 points, with a p-value of 0.09. We conclude that the primary endpoint of the pivotal study does not demonstrate device effectiveness and does not support the benefit of the neuroAD.

Based on a post-hoc analysis of the pivotal study, the sponsor has proposed to limit the patient population to those that have a baseline ADAS-Cog≤30. Because of the post hoc nature of the analysis and the associated uncertainty, FDA typically requests supplemental evidence in an independent dataset to demonstrate the validity of the post hoc findings. In this case, it does not appear that other studies have validated that the ADAS-Cog≤30 is a reliable subgroup that has a higher probability of experiencing benefit than the overall cohort. However, the sponsor is seeking marketing authorization for this subgroup. Restricting the population to those subjects with a baseline ADAS-Cog ≤30, the sham group still outperforms the active group at 7 weeks (+0.47). It is not until the 12-week timepoint (6 weeks after

the treatment ended) that the active group experiences a higher improvement than the sham group (-1.61 points).

We also remind the panel that the ADAS-Cog scale has a total of 70 points and that the changes discussed within this document are very small within the context of the full scale. For this reason, we believe that it is useful to view the data for the group means on the full scale in addition to considering the results in the context of the mean *difference* between the groups from baseline to each assessment timepoint. To clarify the mean changes per group on the full ADAS-Cog scale we provide Figure 44 below which displays the data as the total mean results for active and sham at baseline, 7 weeks (primary endpoint), and 12 weeks. We provide this for the baseline ADAS-Cog≤30 subgroup as the "best case" result and the sponsor's proposed indicated population.

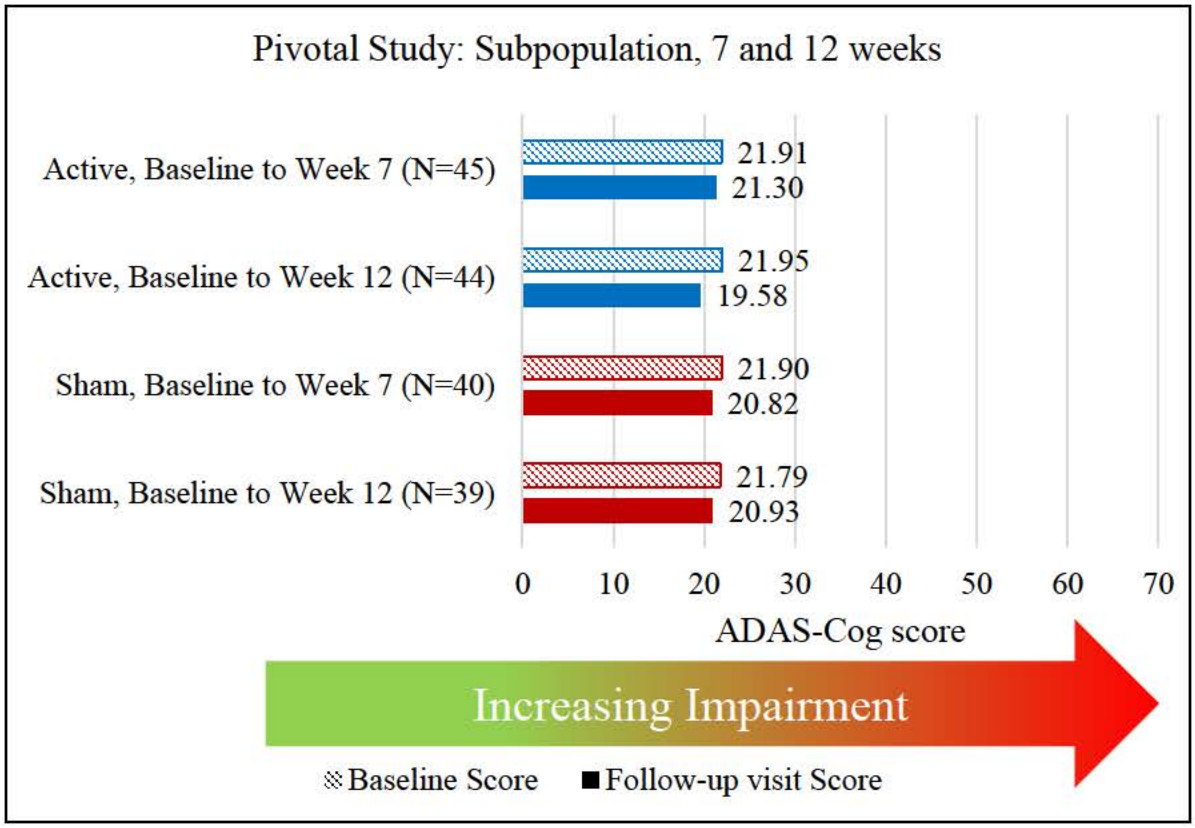


Figure 44.Pivotal Study Post-Hoc Analysis Results, Subgroup, ADAS-Cog Group Means from Baseline to 7 Weeks and Baseline to 12 Weeks

In addition to the pre-specified analysis of the pivotal study and the post-hoc analysis focused on the ADAS-Cog ≤30 subgroup that is discussed above, Neuronix also presented post-hoc analyses using the pivotal study results and an independent Korea Pilot Study in the original submission (Lee, Choi, Oh, Sohn, & Lee, 2016). While the Korea Pilot study demonstrated favorable results for the active group compared to the sham, the subgroup identified from the post hoc analyses of the pivotal study (ADAS-Cog≤30) actually demonstrated a smaller average treatment effect than did the overall cohort. In addition, unlike the US pivotal study, the maximum differences are noted at 7 weeks and not 12 weeks. Although the sponsor provided information from a small, Korean pivotal study of patients with ADAS-Cog<=30, it consisted of only 11 treated subjects.

We have also considered the results of these pooled studies presented by the sponsor (see Appendix V. Post Hoc Analyses Using Supplemental Clinical Data Sources). While the pooled results show a mean change that is more favorable to the neuroAD than the pivotal study at both timepoints in the overall cohort and the subgroup, the trends between 7 weeks and 12 weeks and the results of the ADAS-Cog≤30 are also inconsistent with the pivotal study results.

These inconsistent results (variable subgroup treatment effect, variable time course for treatment effect) raise questions about whether the observations are due to chance, rather than representing a true effect of the device treatment.

Because of these shortcomings and inconsistent findings, there is uncertainty from the available evidence that the device is effective and provides benefits to patients.

Clinically Meaningful Effect

In the absence of a statistically significant improvement in the active group over the sham group in the only prespecified statistical analysis (i.e., the pivotal study primary endpoint) and in the absence of prespecified statistical analysis plans to detect statistically significant differences between groups in a controlled way for the other results provided (e.g., the post-hoc analysis of the pivotal study, the supplemental investigations), the review team has considered whether the results of these analyses are clinically meaningful. We gathered evidence regarding a minimum clinically meaningful difference (MCID) on the ADAS-Cog scale to understand how the results may be translated to clinical practice and whether the numerical changes captured by the scale will be meaningful to patients. We provide a summary of our assessments which include a review of the literature, statements on the Neuronix comparison to approved AD pharmaceuticals, results of Neuronix's patient survey, and discussion points from FDA's Network of Experts.

Review of Scientific Literature

FDA understands that there is no definitive minimum clinically important difference (MCID) on the ADAS-Cog established in the scientific literature. However, when analyzing results of a clinical dataset in which the primary endpoint favored the sham group we look towards the literature to gauge the potential clinical meaning of the post-hoc results of the ADAS-Cog assessment and to give context to the results in the absence of statistical significance. Below is a brief summary of the literature search attempting to uncover a consensus for an MCID on the ADAS-Cog.

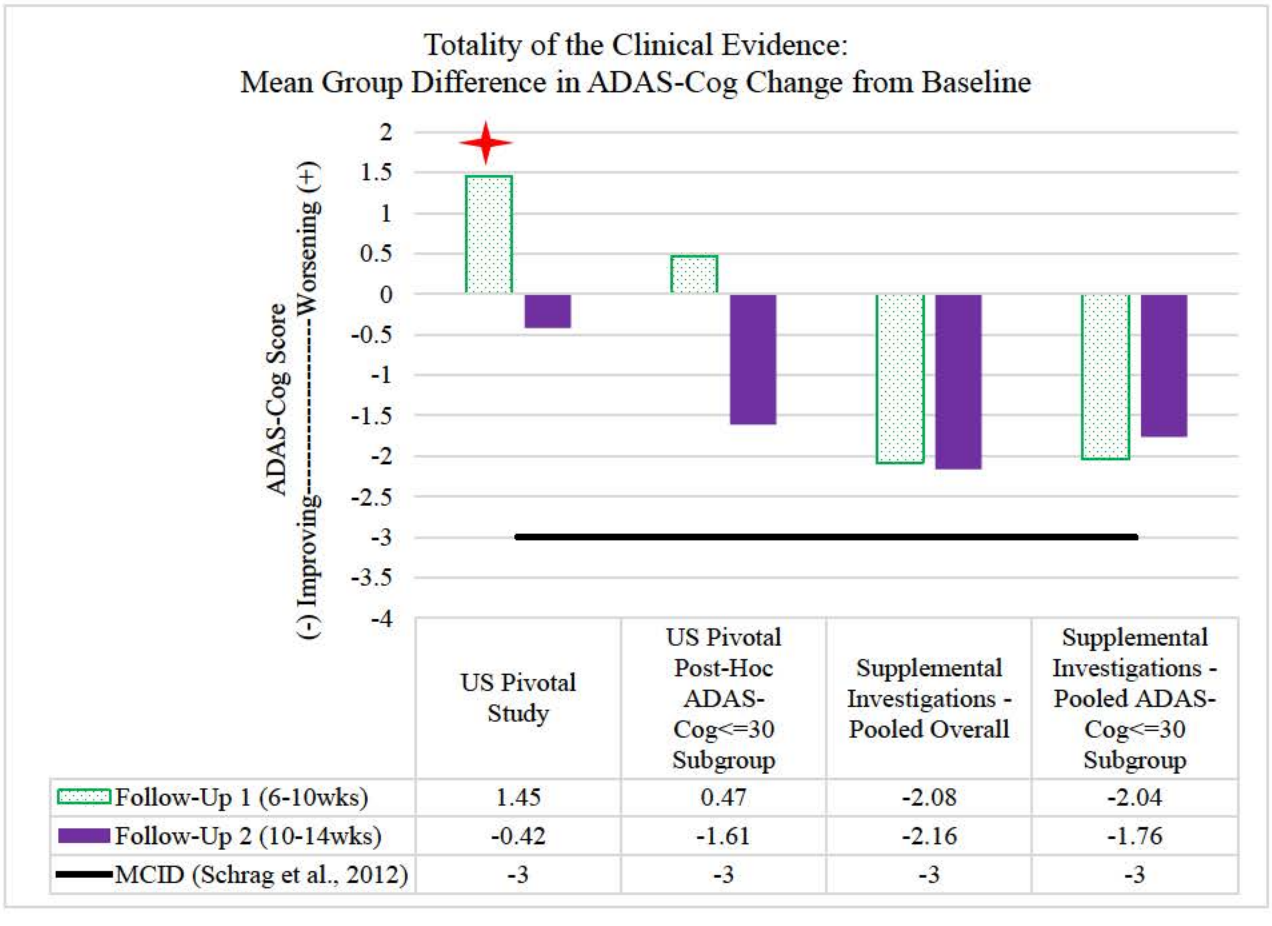
- (Schrag, Schott, & Alzheimer's Disease Neuroimaging, 2012)
  The full text of the Alzheimer's Disease Neuroimaging Initiative (ADNI) study indicates that a 3 point change may be an acceptable MCID for the ADAS-Cog scale for early AD patients. The ADNI study was conducted in an early AD population, and included one hundred and eighty-one patients with baseline ADAS-Cog score=18.5±6.4. The ADNI study population is within the post hoc baseline ADAS-Cog score ≤ 30 population proposed in the IFU for DEN160053. The ADNI study found that the mean ADAS-Cog score changes in patients with clinician judged clinically relevant worsening were 3.1–3.8 points. The mean changes in the clinically unchanged group were 1.9–2.0 points, with the upper 95% CI not exceeding 3 points. Those who deteriorated by one stage on the Clinical Dementia Rating-global scale had a mean change score of 3.98 on the ADAS-Cog scale.

- (Molnar, Man-Son-Hing, & Fergusson, 2009)
    - This article is a systematic review and includes references for the MCID on the ADAS-Cog scale used in drug trials for dementia. MCID measures used in these trials were largely based on expert consensus rather than empirically derived thresholds. The most commonly cited measures of clinical significance was a 4-point change in the ADAS-Cog scale as recommended by a U.S. Food and Drug Administration committee (n=7).

- (Wilcock, Lilienfeld, & Gaens, 2000)
    - In a non-systematic review of the literature, the lowest ADAS-Cog score change identified that was indicated as being clinically meaningful was 2.75 points, due to expected 6 month deterioration* in patients in the placebo group and the magnitude of treatment differences.

      *Please note in this case we have only 12-week data which is not supportive of a disease-modifying/lack of deterioration.

With this background on the clinical meaning and importance of ADAS-Cog scale changes in mind, a summary of the all the studies that were presented in support of the benefit of the neuroAD is provided in the figure below.



	US Pivotal Study	US Pivotal Post-Hoc ADAS-Cog<=30 Subgroup	Supplemental Investigations - Pooled Overall	Supplemental Investigations - Pooled ADAS-Cog<=30 Subgroup
Follow-Up 1 (6-10wks)	1.45	0.47	-2.08	-2.04
Follow-Up 2 (10-14wks)	-0.42	-1.61	-2.16	-1.76
MCID (Schrag et al., 2012)	-3	-3	-3	-3

Figure 45. Applying MCID to Mean Group Differences in ADAS-Cog Change from Baseline

Pre-specified Primary Endpoint of Pivotal Study

Note: In the figure above an example MCID threshold on the ADAS-Cog is drawn at 3-points. This is based on the 3-point MCID reported in (Schrag et al., 2012) and discussion above, and is consistent with the feedback given to Neuronix throughout the review of their submission. In general, these analyses do not appear to demonstrate that device treatment results in clinically meaningful effect compared to sham.

Comparison to Approved AD Pharmaceuticals

In general, there is no legal or regulatory requirement that a device meet or exceed regulatory standards applied to FDA's evaluation of drugs. As such, there is no need for the neuroAD to meet the CDER recommendations that are listed below. However, CDRH's Benefit-Risk Guidance indicates that device benefit-risk decisions should be made in the context of other available treatments. In the sponsor Executive Summary, this comparison between the neuroAD and approved drugs with respect to determining an MCID on the ADAS-Cog scale is presented for the panel to consider. CDRH consulted CDER to provide appropriate feedback to the company regarding the regulatory approval of drugs. The following information is intended to provide the panel with additional background regarding the approval standards of pharmaceuticals:

- CDER recommends that the efficacy of a drug or biologic proposed for the treatment of dementia of the Alzheimer's type (i.e., Alzheimer's disease with overt dementia) be established in an adequate and well-controlled clinical trial; in that trial, evidence of efficacy should have been demonstrated separately (at a p-value ≤ 0.05) on both a cognitive co-primary efficacy measure and a functional (or global) co-primary efficacy measure; the use of a functional or global co-primary efficacy measure is to confirm that the effect on the cognitive instrument be clinically meaningful.

- CDER does not require that the effect of that drug or biologic on either outcome be of a specific (effect) size (as evidenced by some approved drugs for the treatment of AD displaying a difference in ADAS-Cog from baseline of less than 3 points at follow-up (ranging from -1.49 to -2.37)). However, CDER also does not recognize an effect size ≥ 3.0 points (or any other effect size) on the 11-item (standard) Alzheimer's Disease Assessment Scale – Cognitive (ADAS-Cog) as being intrinsically clinically meaningful, and has not approved a drug based solely on an ADAS-Cog endpoint.

- Under all but extraordinary circumstances, CDER recommends that the efficacy of a drug or biologic in dementia of the Alzheimer's type when established as above in a single adequate and well-controlled clinical trial, be replicated in at least one further adequate and well-controlled clinical trial.

- The three acetylcholinesterase inhibitors currently marketed in the United States for the treatment of Alzheimer's Disease are donepezil, rivastigmine, and galantamine. Each was initially approved for the treatment of mild to moderate dementia of the Alzheimer's type based on each showing a statistically significant superiority to placebo in at least 2 adequate and well-controlled clinical trials on a co-primary cognitive endpoint (the ADAS-Cog) and on a co-primary global endpoint, and using a prespecified primary efficacy analysis plan.

Furthermore, the sponsor notes that CDER recently issued a draft guidance document for early stage Alzheimer's disease proposing a single-endpoint threshold for approval of new drugs for early Alzheimer's disease. However, the draft guidance document specifies that to be acceptable as a single primary efficacy outcome measure, the assessment should be "[a]n integrated scale that adequately and meaningfully assesses both daily function and cognitive effects in early AD patients." It is also noted that FDA continues to recommend and accept the independent assessment of daily function and cognitive

effects and cautions that an effect on neuropsychological performance of uncertain independent clinical meaning (e.g., a word recall test) should not allow for an overall finding of efficacy in the absence of meaningful functional benefit.

Sponsor-Provided Physician Survey

Please see Appendix VI. Stakeholder Input for more information on the Physician Survey. Below we provide a summary of the results with respect to the definition of an MCID on the ADAS-Cog scale.

Neuronix concluded the following based on the survey results: "Nearly half of physicians consider at least a 1 point improvement (or less, so long as there is no deterioration) in ADAS-Cog Score clinically meaningful following 3 months of treatment; even more find this threshold clinically meaningful when there is also a 0.5 point improvement in ADCS-CGI-C."

Based on the Physician Survey results, FDA concludes that even when considering an adjunct therapy more than half of physicians considered at least 2 points or greater on the ADAS-Cog score to be clinically meaningful following 3 months of treatment.

CDRH-led Network of Experts

To better understand the clinical meaningfulness in ADAS-Cog score as well as understand how to determine potential patients for Alzheimer's disease interventions and tools available to monitor treatment of cognitive impairment and Alzheimer's disease, the review team consulted CDRH's Network of Experts. Three experts were contacted independently by phone in March-April of 2018 and asked the following four questions:

1. Assuming negligible risk from an intervention, what is the smallest demonstrated change in the ADAS-Cog that you would consider clinically meaningful enough to try the intervention on a patient with Alzheimer's disease?

2. How, if at all, would a patient's baseline score affect the magnitude of the change you would consider clinically meaningful?

3. What instruments do you use in the clinic to track progression of mild cognitive impairment and Alzheimer's disease?

4. If there is a new intervention to treat Alzheimer's disease, what clinical data and information would you use to identify potential patient candidates?

When responding to the question regarding the smallest demonstrated change on the ADAS-Cog that would be considered clinically meaningful enough to try an intervention (assuming negligible risk from an intervention), two out of the three experts stated that a 4 or 4-5 point change on the ADAS-Cog was clinically meaningful. The third expert noted reading in literature a 2-3 point change as being clinically meaningful for a disease-modifying therapy; however, the studies submitted to FDA were neither designed to target a disease-modifying therapy nor evaluated to assess whether the course of disease changed.

When asked about whether a patient's baseline score would affect the magnitude of change on the ADAS-Cog that would be considered clinically meaningful, two experts hypothesized that you may need a larger change from baseline in mild patients than in moderate to advanced patients because with moderate to advanced patients you are expecting rapid decline; a smaller change in advanced patients may be

clinically meaningful. A third expert was not able to answer the question because they felt that the ADAS-Cog is not a useful test in moderate and severe disease.

FDA Conclusion of Benefit

Based on the resources discussed above, a range of 2-5 point change on the ADAS-Cog scale has been suggested as a minimum clinically important difference (MCID).

Based on the available clinical data analyses (pivotal study, pivotal study post-hoc analyses, pooled Korea and pivotal analyses, supplemental pooled analyses), the supplemental pooled "overall" dataset is the only dataset that supports the lowest value of 2-points at both follow-ups. The supplemental pooled results do not include the pivotal study. The supplemental pooled data limited to the ADAS-Cog≤30 subgroup reaches two points only at the first follow-up (6-10 weeks).

Regarding the secondary assessment, the CGI-C for this population remains consistently around a score of 4 (no clinical change). This does not appear to support clinical benefit as a companion to the ADAS-Cog data.

In summary, because the pivotal clinical study failed to meet its prespecified primary endpoint, the high degree of uncertainty in the pooled supplemental analysis, the uncertainty in the post hoc pivotal study analysis, and the failure to consistently demonstrate clinically meaningful results when compared to sham, it is difficult to conclude that the available data demonstrates effectiveness or a clinical benefit of the neuroAD as an adjunctive treatment in the intended population.

The FDA requests panel input on whether the device provides a benefit to patients in the intended population, based on the available information.

FDA Summary of neuroAD Risks

Probable Risks Noted from Pivotal Study Results

The rate of pivotal study subjects experiencing any definite, probable or possible study procedure or device-related adverse event was 14% (11/79, includes the 20 non-randomized active patients) in the active group and 4% (2/50) in the sham group. The 11 active group subjects reported 15 events total that were found to correlate with relatedness to the device or procedure. These AEs were reported and determined by the site investigator. The AEs were described to be mild per the CSR and among the expected AEs previously reported with TMS. The AEs included headache, neck pain, skin discomfort or muscle twitching. The CSR states that some of the AEs persisted through multiple treatment sessions which were described as mild, transient, not requiring treatment discontinuation, occurring during treatment, and were managed by adjusting the treatment intensity (MT%).

All adverse events related to the study device resolved without sequalae.

Probable Risks Noted from Supplemental Investigations

As discussed in the Appendices, the supplemental investigations do not contain complete and detailed reporting of Adverse Events which may be of concern. Based on the patient-level raw data that FDA has available for these studies, we note that the Korea Pilot reported there were no AEs (25/27 patients raw data received). In the Korea Pivotal (Korea -2) study there was one reported AE in the 22 patients (11 active) which was a skin rash attributed to an allergy. The Italy study (13 patients, 6 active) reported no

AEs of any type. No raw AE data from the Assaf-1, Assaf-2, and Assaf-3 trials is available to FDA. This would account for the primary difference between the sponsor's AE reporting and the FDA when discussing the Supplemental Investigations.

For completeness, we note that the sponsor did report descriptions of AEs that occurred in some of these supplemental studies. Because these appear to be potentially related AEs only, and we do not know the safety or monitoring plans of these studies, these descriptions carry uncertainty. However, in order to characterize the potential risks of this device, the potentially-related AEs listed for the active patients are as follows: psychiatric symptoms that required medication (n=1, deemed unrelated to the device per the PI), mild and transient hearing impairment post-intervention, blurry vision (potentially from computer screen), eye pain, neck pain/stiffness, mild scalp pain, soreness at stimulation site, achiness, fatigue, nausea, transient eye heaviness, mild to moderate headache events, tiredness, dizziness, increased anxiety.

FDA Conclusion of Probable Risk

Within the collection of data provided, there were no device-related serious adverse events but there was a higher number of adverse events in the treatment group compared to the sham. Overall, the risk appears to be low. The FDA does not have significant safety concerns with this device based on the risk profile that has been presented.

The US Pivotal study reported that 14% of the active participants reported at least one AE *related* (possibly, probably, definitely) to the device or procedure. It is assumed that the AEs being reported in the Supplemental Investigations are also device or procedure related only and none of the other health related AEs were reported. If this is not the case, the rate of AEs in these studies appears to be much lower than anticipated for a population primarily over age 65.

It should be noted that the safety data that FDA has analyzed is limited to that collected within the prespecified timeframe of the studies which extends to a maximum of 14 weeks after the final treatment session. Because AD patients are a relatively new population for the use of TMS, as of 2018, there have been few published longer term (six or more months) and larger sample size studies to assess AEs that may be relevant to this specific population. (Chang, Lane, & Lin, 2018)

Conclusion

Advancing therapies for Alzheimer's disease is important to public health. At the same time, patients and their families are best served when they can understand the benefits and risks of marketed products and rely on these products to provide a reasonable assurance of safety and effectiveness in a predictable and repeatable manner. In the case of the neuroAD device, while it is a low risk device, it is difficult to conclude that the available data demonstrates effectiveness or a clinical benefit. As a result, it is difficult to conclude that the benefits of the device outweigh the risks.

FDA is seeking input from the panel on interpretation of: 1) the primary pre-specified endpoint of the pivotal study (ADAS-Cog at 7-weeks) that demonstrated non-statistically significant improvements in the sham group compared to the neuroAD treatment group (sham vs. treatment, -1.38 points vs. 0.07 points, p=0.09); 2) the pre-specified secondary endpoint assessments of the pivotal study (ADAS-Cog at 12 weeks after the treatment ended at 6 weeks (sham vs. treatment, -0.61 points vs. -1.03 points, p=0.64) and CGI-C at 7 weeks (sham vs. treatment, 4.06 points vs. 4.04 points, p=0.96) and 12 weeks (sham vs. treatment, 4.19 points vs. 3.84 points, p=0.12); 3) post-hoc analysis of the pivotal study (see Pivotal Clinical Study – Post-Hoc Analysis of Baseline ADAS-Cog≤30 Subgroup); and 4) analyses of supplemental datasets (see APPENDICES II through IV).

Specifically, the FDA has significant concerns with the clinical data presented because the prespecified analysis of the pivotal study results favored the sham group over the treatment group and the post hoc analyses assessing device effectiveness carry significant uncertainty. There is uncertainty in the results of the post-hoc analysis of the pivotal study because the post-hoc analysis was conducted after the results of the trial were known, introducing bias. In the supplemental analyses there is uncertainty with respect to the conditions of study conduct (e.g., data collection, reporting and documenting of adverse events (some missing)) and the analysis methods (e.g., data pooling methodology, lack of pre-specified metrics to limit bias). In addition, the supplemental data and analyses presented show inconsistent results raising concerns that the observations are due to chance, rather than a true effect of the device treatment.

Importantly, as previously mentioned, the prespecified analysis of the pivotal trial did not demonstrate a clinically or statistically meaningful benefit. We also note for the following concerns with the pivotal study results in more detail for the panel:

- FDA could not identify any factors to support the contention that the 7-week results in the pivotal study do not provide a good estimate of the treatment effect.
- FDA could identify no plausible explanation for the differences seen between 7 weeks and 12 and cannot confidently attribute them to a delayed device treatment effect in the absence of any intervention occurring between 7 weeks and 12 weeks.
- FDA is unable to explain why the "recovery" in the active group at 12 weeks is more pronounced in patients who are within the proposed baseline ADAS-Cog≤30 population.
- The post-hoc analysis of the pivotal study found the best case result in favor of the device was a -1.61 improvement over sham in the ADAS-Cog <=30 subpopulation at the secondary endpoint of 12 weeks. We have concerns with both the reproducibility of the post-hoc subgroup analysis of the pivotal trial and the clinical meaningfulness of the post-hoc result.

Though the pivotal study did not meet its primary endpoint and the post hoc analysis results in high levels of uncertainty, we also considered the additional datasets provided by Neuronix. Neuronix provided new studies and additional analyses on the ADAS-Cog endpoint and the CGI-C endpoint. This included data from 118 new patients (97 active, 21 sham) from various small studies and treatment clinics. Neuronix pooled these supplemental datasets (which did not include the pivotal study) and reported change in ADAS-Cog results for the 6-10 week from baseline range and the 10-14 week from baseline and CGI-C

results for the 10-14 week from baseline timeframe. The best case change on the ADAS-Cog using this updated analysis was an improvement of -2.16 over sham at the 10-14 week range in the overall cohort (not limited to the intended population of baseline ADAS-Cog≤30). As described in Appendix V. Post Hoc Analyses Using Supplemental Clinical Data Sources, the general trend in the supplemental studies is for the therapeutic gain to decrease over time and for the subgroup to perform worse than the overall cohort (in contrast to the pivotal study).

In total, the sponsor provided a large clinical data set with primary results that did not demonstrate effectiveness. The ancillary analyses and supplement studies present conflicting results and are difficult to interpret. In reviewing the totality of the evidence, it is difficult to conclude that the available data demonstrates effectiveness or a clinical benefit of the neuroAD in the intended population. As a result, it is difficult to conclude that the benefits of the device outweigh the low risks.

FDA is committed to fostering the development of products that can meet unmet clinical needs. The neuroAD was granted breakthrough status because the device, if demonstrated to be safe and effective, would treat Alzheimer's disease, addressing an unmet clinical need. The FDA Benefit-Risk Guidance Document provides clarification regarding the benefit-risk principles as applied to a device that is intended to address an unmet medical need (U.S. Food and Drug Administration, 2016b):

> "A device may address unmet medical need by providing a clinically meaningful advantage over existing technologies, providing a greater clinically meaningful benefit than existing therapy, posing less risk than existing therapy, or providing a treatment or means of diagnosis where no alternative is available. It is not unusual for novel devices that address an unmet medical need to have relatively small probable benefits, and FDA may determine the novel device to be reasonably safe and effective even though the applicant demonstrates a relatively small probable benefit. […] In these circumstances, in order to facilitate patient access to new devices important for public health and to encourage innovation, we may tolerate greater uncertainty in an assessment of benefit or risk than for most established technologies, particularly when providers and patients have limited alternatives available."

FDA acknowledges that the population of patients diagnosed with Alzheimer's Disease is underserved by current device options, and that the neuroAD system would be the first device intervention to treat the symptoms of Alzheimer's Disease if approved. In this document, we have provided a summary of the available information and our perspective that the available evidence is not sufficient to establish that the device provides a clinically meaningful effect or that the benefits outweigh the risks.

Because of the importance of this patient population and our desire to bring novel treatments to patients, and because we are committed to an open and transparent process, we are seeking panel input on the assessment of benefits and risks of the device. While we have shared our current thinking and summarized our assessment, FDA is seeking panel input before rendering a final decision on the submission as to whether the information provided demonstrates a reasonable assurance of safety and effectiveness as defined in 21 CFR 860.7(d)(1) and (e)(1). Summarized, the evidence must show that when using the device properly, the probable benefits to health outweigh the probable risks and there is an absence of unreasonable risk (safety), and that there are clinically significant results in a significant portion of the target population (effectiveness).

Today's panel is intended to further explore the data submitted to support its use for Alzheimer's disease and to discuss the benefit-risk ratio for the neuroAD.

Panel Question: The Panel will be asked to discuss and make recommendations on whether the probable benefits to health outweigh the probable risks.

APPENDICES

Appendix I. Nonclinical Studies

All questions regarding the non-clinical testing have been addressed to FDA's satisfaction. The following sections describe the testing and evaluations completed by the company.

Sterilization/Shelf Life/Reuse

The neuroAD Therapy System Operator's Manual contains cleaning instructions. The system is not sold sterile, is not intended to be sterilized by the user, and has no stated shelf life.

Biocompatibility

The only patient-contacting components of the neuroAD are the Patient Cap and the Patient Head Marker band. Both components have limited duration contact with skin (surface contacting, less than 24 hours).

- The Patient Cap is an off-the-shelf, repeated use (for a single patient), medical grade product made from cotton and manufactured by Scrubs.com (Burlingame, CA, USA).

- The Patient Head Marker band is placed around the patient's forehead and is used to secure the Patient Head Marker. The Patient Head Marker band is made of VELSTRETCH® Brand Loop 151 Fastener, manufactured by Velcro USA Inc. A certification of the biocompatibility of the product was provided in the original de novo submission.

The magnetic coil is not patient contacting due to the barrier created by the Patient Cap but the coil is biocompatible. If the neuroAD was used without the Patient Cap, the magnetic coil contact would be limited duration with the skin (surface contacting, less than 24 hour duration). The magnetic coil of the neuroAD Therapy System is identical to the previously cleared Magstim 3530-00 70mm Double Rapid Air Coil (K051864).

Software Testing

The neuroAD uses software to deliver, monitor, and adjust the TMS therapy and cognitive training and to facilitate accurate positioning of the coil via the navigation system. The Base Unit and the Navigation Unit each contain designated software. The neuroAD does not contain any networked communications.

The Base Unit PC software is responsible for the following functions: (a) start treatment, (b) stop treatment, (c) display paradigm, (d) save paradigm data to USB, (e) perform training, and (f) gather patient information.

The Base Unit DSP software is responsible for generating and monitoring the TMS therapy which includes the following functions: (a) charge capacitors, (b) discharge capacitors, (c) generate train of pulses, and (d) monitor temperature and voltages.

The Navigation PC software is responsible for performing coil calibration, patient registration, brain target selection, and navigation (including providing user instructions

As a system, the neuroAD software represents a MODERATE level of concern based on the FDA Guidance Document, "Guidance for the Content of Premarket Submissions for Software Contained in

Medical Devices"(U.S. Food and Drug Administration, 2005). The sponsor has provided the appropriate documentation and testing for this level of concern.

Note: The pivotal study was initiated with Base Unit PC software version 1.1.15 but during the study a software error related to saving patient data after the backup procedure was detected and the software was updated to version 1.1.16. The software update had no impact on the TMS therapy but did impact the cognitive training level of some of the cognitive paradigms. Only two patients were identified as being affected by the software error and were excluded from the statistical analysis of the pivotal study.

Electrical, Mechanical, and Thermal Safety

The neuroAD has been tested for electrical, mechanical, and thermal safety according to IEC 60601-1:2005(R)2012. This included the cooling test and the acoustic test. The Base Unit and the Navigation Unit were tested separately. While IEC 60601-1 recommends testing as a system, because of the way the devices function, testing separately is acceptable. Both the Base Unit and Navigation Unit passed all electrical, mechanical, thermal, and acoustic tests.

Electromagnetic Compatibility (EMC)

As a device that may susceptible to electromagnetic interference and may cause interference, the neuroAD has been tested for electromagnetic compatibility according to IEC 60601-1:2:2007(R)2010. The Base Unit and Navigation Unit were not tested for EMC as a system but were tested separately. While IEC 60601-1-2 recommends testing as a system, because of the way the devices function, testing separately is acceptable. After a minor design change to the Base Unit which is reflected in the marketed product, the Base Unit and the Navigation unit passed all applicable EMC tests.

Design Verification and Validation Testing

In addition to the system testing described above, the following design verification and validation testing was also presented to FDA. The device met all acceptance criteria.

Table 20. Design Verification and Validation Testing: Base Unit

Bench Test Type	Applicable Standard or Guidance
Magnetic Field Spatial Distribution Measurements	FDA Guidelines for rTMS: Guidance for Industry and Food and Drug Administration Staff - Class II Special Controls Guidance Document: Repetitive Transcranial Magnetic Stimulation (rTMS) Systems (Document issued on: July 26, 2011)
Magnetic Field Strength Gradient Measurements	FDA Guidelines for rTMS: Guidance for Industry and Food and Drug Administration Staff -Class II Special Controls Guidance Document: Repetitive Transcranial Magnetic Stimulation (rTMS) Systems (Document issued on: July 26, 2011)
Cooling Test	IEC 60601-1: 2005 + CORR. 1:2006 + CORR. 2:2007 + AM1:2012 - Medical electrical equipment-Part 1: General requirements for basic safety and essential performance
Acoustic Test	IEC 60601-1: 2005 + CORR. 1:2006 + CORR. 2:2007 + AM1:2012 - Medical electrical equipment-Part 1: General requirements for basic safety and essential performance

Table 21. Design Verification and Validation Testing: Navigation Unit

Bench Test Type	Applicable Standard or Guidance
Optical Navigation Accuracy	ASTM, F2554-10 - Standard Practice for Measurement of Positional Accuracy of Computer Assisted Surgical Systems FDA Guidelines for TMS: Class II Special Controls Guidance Document: Repetitive Transcranial Magnetic Stimulation (rTMS) Systems (Document issued on: July 26, 2011)
Registration and Calibration Accuracy	FDA Draft Guidance: Applying Human Factors and Usability Engineering to Optimize Medical Device Design FDA Guidelines for TMS: Class II Special Controls Guidance Document: Repetitive Transcranial Magnetic Stimulation (rTMS) Systems (Document issued on: July 26, 2011)
Photobiological safety of lamps and lamp systems	IEC62471 - Photobiological safety of lamps and lamp systems

The sponsor also conducted a Usability Study on the overall neuroAD Therapy System according to the FDA Draft Guidance, "Applying Human Factors and Usability Engineering to Optimize Medical Device Design" This Guidance Document was finalized on February 3, 2016 (U.S. Food and Drug Administration, 2016a).

Appendix II. Supplemental Investigations Study Designs

The following provides study design narratives for all the investigations that were used in the post-hoc analyses for the neuroAD.

Maintenance Therapy
In two of the supplemental clinical datasets (Assaf-1 and Assaf-2) Maintenance Therapy was used. Maintenance Therapy consisted of an additional 3 months of treatment with 2 sessions per week following the standard treatment protocol of 6 weeks, 5 days a week. Each Maintenance Therapy session included 1300 pulses for 3 alternate brain regions. These studies did not have observations in the 10-14 week timeframe so they were not included in the ADAS-Cog analysis at the 10-14 week timeframe. However, the decision of the Assaf investigators to include a maintenance therapy protocol is noted here for panel consideration given the FDA uncertainty in the reliance of a 12 week treatment effect (7 weeks post-treatment end).

Supplemental Investigations Study Designs

A summary of the study designs used for the Supplemental Investigations is provided below. In some cases, the ADAS-Cog was not used, or was not collected at 7wks and 12wks (to enable comparison with the pivotal study.) Therefore, in the pooled assessment, Neuronix has provided ADAS-Cog results within two timeframes of 6-10 weeks and 10-14 weeks. In the case of no ADAS-Cog collection, the data were not pooled.

Randomized Controlled Trials

Assaf-2

Assaf-2 was provided in S002 as a Supplemental Investigation. It is also listed in the original submission as Pilot Study 2 (Rabey et al., 2013). This study included Maintenance Therapy.

The study employed the former system configuration, the NICE-XP1 (please see Device Description section above.) It was a randomized, double-blind, single site study conducted in Israel. Assaf-2 is a second pilot study that was conducted at Assaf Harofe Medical Center, Israel (the first study conducted at this site is listed as Assaf-1 below in the open-label section and also as Pilot Study 1 in the original submission, (Bentwich et al., 2011)

Principal eligibility criteria included mild-moderate AD patients, age 55-85, MMSE in the rage 18- 24, with a CDR score of 1. A total of 15 patients were enrolled in the study, of which 7 were randomized to the treatment arm and 8 to the placebo arm. All were previously diagnosed AD patients. There were no statistically significant differences between the groups, on any of the baseline characteristics. 13 patients were medicated, and 2 patients were non-medicated. 12 patients had mild AD, and 3 had moderate AD.

The study protocol provided for 6 weeks of intervention (6w), followed by 12 weeks of 1-2 treatments per week as a **maintenance phase** (4.5m). The primary endpoints were safety and the ADAS-Cog at 7 weeks and 18 weeks compared to baseline, relative to respective change in placebo. Secondary endpoints included the CGI-C and neuropsychiatric inventory (NPI).

Assaf-3

Assaf-3 was provided in S002 as a Supplemental Investigation. It is also listed in the original submission as Pilot Study 4. This study has not been published.

The study employed the proposed neuroAD system configuration. It was a randomized (2:1), double-blind, single site study conducted in Israel. Assaf-3 is a third pilot study that was conducted at Assaf Harofe Medical Center, Israel (the first study conducted at this site is listed as Assaf-1 below in the open-label section and also as Pilot Study 1 in the original submission, (Bentwich et al., 2011); the second is Assaf-2 described above)

Eligibility criteria were the same as in the prior studies, i.e., mild-moderate AD patients, MMSE in the range of 18-24, and a CDR score of 1. A total of 16 patients were recruited, of which 10 were randomized to the treatment arm and 6 to the placebo arm (2:1 randomization). All were previously diagnosed AD patients. There were no statistically significant differences between the groups on any of the baseline characteristics.

The protocol provided for 6 weeks intervention (6w), with 6 weeks follow-up, without any further treatments (12w). The primary endpoints were safety and the ADAS-Cog improvement at 7 weeks compared to baseline relative to respective change in placebo. The secondary endpoints included the CGI-C and MMSE.

Beth-Israel (Harvard)

Beth-Israel/Harvard was provided in S002 as a Supplemental Investigation. It is also listed in the original submission as Pilot Study 3. It does not appear to be published but under peer review as of May 2015 ("Combined brain stimulation and cognitive training in Alzheimer's disease, Brem A., et al.).

The study employed the former system configuration, the NICE-XP1 (please see Device Description section above.). It is a single center, double-blind, randomized, placebo-controlled study conducted at Beth Israel Deaconess Medical Center, Boston, MA, USA.

The study involved 3-arms: treatment and placebo arms and a 3rd arm comprised of real cognitive training (as in the Treatment Group), but sham TMS. This 3rd arm was added in order to investigate the patients' performance with cognitive training alone rather than the combination of cognitive training with TMS.

As in the previous studies, eligibility was limited to mild-moderate AD patients, with MMSE in the range of 18-24, and a CDR score of 1. A total of 21 patients were recruited, divided into 3-arms: 10 in the treatment arm, 6 in the placebo arm, and 5 in the cognitive training arm. The study was conducted in two phases, treatment vs. placebo, with 5 and 6 subjects enrolled, respectively; and treatment versus cognitive + sham TMS, with 5 and 5 subjects enrolled, respectively. There were no statistically significant differences between the groups on any of the baseline characteristics

The protocol provided for 6 weeks intervention (6w), with 4 weeks follow-up, without any further treatments (10w). The primary endpoints were safety and the ADAS-Cog at week 7 and week 10 compared with baseline among the three arms. The secondary endpoints included the CGI-C, ADAS-ADL, GDS, and various neuro-physiological and plasticity tests.

Korea Pilot (Korea-1)

The Korea Pilot study was provided in S002 as a Supplemental Investigation. It is also listed in the original submission as Pilot Study 5 (Lee et al., 2016) and was pooled with the pivotal study as part of a post-hoc analysis (Appendix III. Post-Hoc Analyses Combining Korea Pilot Study with U.S. Pivotal Study)

The study employed the proposed neuroAD system configuration. This is a single center, double-blind, randomized (2:1), placebo-controlled study conducted at Chungnam National University Hospital, Daejeon, Korea. This study was initiated independently by the institution and was not sponsored by Neuronix.

The eligibility criteria were similar to those in prior studies, i.e., mild-moderate AD patients, with MMSE in the range of 18-24, and CDR score of 1 or 2. A total of 27 patients were recruited, of which 18 were randomized to the treatment arm and 9 to the placebo arm (2:1). All were previously diagnosed AD patients. There were no statistically significant differences between the groups, on any of the baseline characteristics.

The protocol provided for 6 weeks intervention (6w), with 6 weeks follow-up, without any further treatments (12w). The primary endpoints were safety and the ADAS-Cog at 7 weeks compared to baseline relative to the change in the placebo group. Secondary endpoints included the CGI-C and NPI.

In Supplement 2, Neuronix included long-term data from 5 active patients and 3 sham patients that returned to the clinic after study conclusion. The source data presented for this new data is a poster entitled "Repetitive transcranial magnetic stimulation combined with cognitive training in Alzheimer's disease" with authors Juyoun Lee, Eungseok Oh, Eun Hee sohn, and Ae Young Lee from Chungnam National University Hospital, Deajeon, South Korea.

Korea Pivotal (Korea-2)

In Supplement 2 Neuronix provided interim data from a new source called the Korea Pivotal study (the study was still in progress at the time of FDA supplement submission). Neuronix states that the Korean Pivotal Study's design is very similar to the neuroAD™ Pivotal Study, except that the Korean Pivotal Study enrolled only mild AD patients with a baseline ADAS-Cog ≤ 30). This study was designed for the propose of supporting a future K-FDA application. The Korea Pivotal study used the proposed neuroAD system configuration.

The eligibility criteria were similar to those in prior studies, i.e., male and female subjects who are diagnosed with mild to moderate AD, with MMSE 21-26, CDR=1, and baseline ADAS-Cog>17. Subjects were randomized into Active and Sham group in 1:1 ratio. The interim data included 22 subjects, 11 active and 11 sham.

Like the Korea Pilot and the pivotal study, the protocol provided for 6 weeks intervention with follow-up at Week 7, and on Week 12. No maintenance treatment sessions performed. At the time of de novo submission, this study was ongoing and interim results for ADAS-Cog change from baseline and between groups were presented for week 7 and week 12. The sponsor has advised that the study was placed on hold based on a determination by the Korean-FDA that it wishes to receive the US FDA's decision prior to continuation of the Korean Pivotal Study and rendering its decision.

Italian

In Supplement 2 Neuronix provided a summary of a new Italian Study. This was a double-blind, randomized, sham-controlled study, conducted at Universita Cattolica del Sacro Cuore Institute of Neurology (Gemelli), Rome, Italy. This study is unpublished. The study employed the proposed neuroAD system configuration.

Eligibility criteria included male and female subjects, age 60-90, who were diagnosed with mild to moderate AD, with MMSE 18-26 and baseline ADAS-Cog>17. Similar to the Harvard Study, the Italian Study involved 3-arms: treatment and sham arms, and a third arm comprised of real cognitive training (as

in the Active group), but sham TMS. A total of 13 subjects were recruited, divided into 3-arms: 6 in the Active arm, 2 in the Sham arm, and 5 in the Cognitive training arm.

The protocol provided for 6 weeks intervention (6w). All subjects were followed on Week 7 (immediately post treatment), and Active treatment subjects were also followed on Week 10.

Neuronix notes and FDA agrees that the results of this study should be interpreted with caution due to the small sample size – Active group – 6 subjects, Sham group – 2 subjects, Cognitive training only group – 5 subjects

Thai

In Supplement 2 Neuronix provided a summary of a new Thai Study. Few details were provided about this study design. Neuronix lists this data as a controlled clinical study. The source data presented for this new data is a poster entitled "Pilot study of rTMS in mild to moderate AD for 30 sessions: effect on cognition and gait performance at 6-months" with authors Vorapun Senanarong, Sunee Bovonsunthonchai, Nuttapol Aoonkaew, Atthapol Raksthaput, and Suthipol Udomphanthurak from Mahidol University, Bangkok, Thailand. The study employed the proposed neuroAD system configuration.

Main entry criterion is listed as probable AD diagnosis. The data includes 9 subjects – 6 active and 3 cognitive training only.

Intervention lasted 6 weeks with follow-up occurring at 6 months. Assessment scales include the TMSE.

Open Label Trials

Assaf-1

Assaf-1 was provided in S002 as a Supplemental Investigation. It is also listed in the original submission as Pilot Study 1 (Bentwich et al., 2011). This study included Maintenance Therapy.

The study employed the former system configuration, the NICE-XP1 (please see Device Description section above.) It was a single-arm single site study conducted in Israel. Assaf-1 is first of three studies conducted at Assaf Harofe Medical Center, Israel (see Assaf-2 and Assaf-3 above).

Principal eligibility criteria included mild-moderate AD patients, age 55-85, MMSE in the range of 18-24, and with CDR score of 1. In total, 8 patients were enrolled, all of whom were previously diagnosed AD patients. Duration from diagnosis was 1.5 – 3 years. Of the 8 patients, 2 patients were non-medicated, 4 patients took ChEI, and 2 patients took ChEI + NMDA. 7 patients were mild (by MMSE and by ADAS≤30); 1 patient was moderate (by MMSE 19, and by ADAS>30).

The protocol provided for 6 weeks of intervention (6w), followed by 12 weeks of 1-2 treatments per week as a **maintenance phase (4.5m)**. The primary endpoints for the study included safety, ADAS-Cog at 7 weeks and at 18 weeks compared to baseline, and CGI-C after 7 weeks and 18 weeks compared to baseline; the secondary endpoints were the MMSE, ADAS-ADL, Hamilton, and NPI.

Commercial Clinic Experience (OUS)

Nantes -France

From Supplement 2, this was an open-label, retrospective, naturalist follow-up of commercial patients treated at University Hospital, Nantes, France. 10 patients received Active treatment and were followed for a period of up to 6 months.

It is important to note that, at baseline, subjects had and ADAS-Cog score ranging from 6.5 to 36

and a MMSE score ranging from 12 to 26. Main eligibility criteria was male and female subjects who are diagnosed with probable AD.

Patients were given the same treatment schedule as previous studies: 6 weeks, 5 days a week

(overall 30 sessions). Subjects were followed up at Week 7 (day 45), and on 6 months. The primary assessment was provided as a change in ADAS-Cog and MMSE from baseline after 6 weeks of treatment.

NeuroCare Clinics-Israel

NeuroCare Clinics-Israel (Cohort 1) was provided in the original submission as Pilot Study 6 and publication (Rabey & Dobronevsky, 2016); Cohort 1 included 30 subjects. 54 additional subjects were included as Cohort 2 in Supplement 2. To date – 80 subjects participated in the program with 84 total treatments recorded (approximately four subjects repeated the treatment course).

Cohort 1

Patients were provided treatment using the proposed neuroAD Therapy System configuration. Patients from this cohort were from two clinics in Israel – in Ramat Gan (Israel) and Haifa (Israel).

As the clinics were offering the treatment as part of commercial program (e.g. – paid for), patients were recruited following advertising campaigns in Israel. All patients who received treatment with the Neuronix device for mild-moderate AD at these centers were asked to participate, of which 30 were enrolled. The protocol provided for 6 weeks intervention (6w). 5 patients came for a follow-up after 10 months (10m) and received a 2nd intervention course (12m). The remaining patients had no additional treatments. The primary endpoints included safety and the ADAS-Cog; the MMSE was a secondary endpoint.

Cohort 2

Cohort 2 provided an additional 54 patients to the findings of Cohort 1. Cohort 2 consists of subjects who were previously diagnosed with Alzheimer's Disease. The data now includes patients from 3 clinics in Israel. Both versions of the device (NICE-XP1 & neuroAD) were used in the program. As of the end of 2014, NICE-XP1 is obsolete.

Treatment course remains 6 weeks, 5 days a week (overall 30 sessions).

Program procedures include administration of ADAS-Cog and MMSE both at baseline and follow-up after treatment. Follow-up varies based on subjects' availability and program changing procedures throughout the years. It is important to note that some of the subjects are coming from abroad, which requires deviations from standard procedures.

Neuronix-UK

Neuronix-UK was provided in S002 as a Supplemental Investigation. Few details regarding Neuronix UK are provided. Data included 4 patients from a naturalistic follow-up after commercial use. Treatment was

6 weeks, 5 days a week and ACE III was administered at Week 7. Patients with probable AD diagnoses were referred to this treatment.

High Wycombe-UK

High Wycomb-UK was provided in S002 as a Supplemental Investigation. Few details regarding this site are provided. Data included 6 patients from a naturalistic follow-up after commercial use. Treatment was 6 weeks, 5 days a week and ACE III was administered at Week 7. Patients with probable AD diagnoses were referred to this treatment.

Orsay – France

Orsay-France was provided in S002 as a Supplemental Investigation. Few details regarding this site are provided. Data included 10 patients from a naturalistic follow-up after commercial use. Treatment was 6 weeks, 5 days a week and MMSE was administered at Week 7. Patients with probable AD diagnoses were referred to this treatment.

FDA Summary Comments on Study Designs of Supplemental Data

The FDA cautions the panel members on the small sample sizes, the limited information regarding the pre-specified statistical plan for each individual study, and the lack of pre-specified pooling plan. Demonstrating the poolability/applicability of the OUS patient populations to the intended US populations is also a consideration that is typically applied to OUS datasets.

In most cases, the studies listed as Supplemental Investigations were conducted independently of Neuronix. However, Neuronix has disclosed information regarding their involvement in some of the studies:

- Neuronix was a co-sponsor for the Italian Study along with the distributor. Contributions paid by Neuronix to the study investigators amounted to approximately $6,500 during the entire study.

- Neuronix sponsored the Assaf 3 Study; however, contributions paid by Neuronix to the study investigators amounted to less than $10,000.

- No investigators currently own stock in Neuronix, however, Dr. Martin Rabey has received stock options and other payments in connection with his work as an advisor to Neuronix. Since 2010, Dr. Rabey received options to purchase 2,440 of Company's Ordinary Shares in Neuronix; these options are not fully vested as of today. Neuronix is not a publicly traded company and therefore shares are not available for public purchase. In addition since 2010, Martin was paid by both Neuronix and its fully owned subsidiary NeuroCare a total sum of approx. $199,004.
    - Dr. Rabey was on the Medical Monitoring Committee of the Pivotal Study. The medical committee provided independent medical support for the study, including review of Serious Adverse Event (SAE) reports, and providing study investigators with support on medical issues and questions. The medical committee also performed blinded review of protocol deviations and adverse events at the end of the study and prior to study unblinding
    - Dr. Rabey was also the principal investigator for the Assaf 1 Study, Assaf 2 Study, and Assaf 3 Study. Dr. Rabey disclosed his information to the ethical committees prior to the studies being conducted.
    - Dr. Rabey currently acts as the clinical director of the NeuroCare clinic in Israel.

Appendix III. Post-Hoc Analyses Combining Korea Pilot Study with U.S. Pivotal Study

In addition to the pre-specified analysis of the pivotal study and the post-hoc analysis focused on the ADAS-Cog ≤30 subgroup that is discussed above, Neuronix also presented post-hoc analyses using the pivotal study results and an independent Korea Study in the original submission (Lee et al., 2016). For all post-hoc analyses, Neuronix presented data on the original primary assessment timepoint of 7 weeks as well as the secondary assessment timepoint of 12 weeks.

Note: The analyses presented in this section are post-hoc analyses without pre-specification and multiplicity adjustment. Therefore, we do not include p-values.

Korea Pilot Study

Neuronix presented data from a 27 subject Korea Pilot study (Lee et al., 2016) pooled with the Pivotal Study data. Neuronix describes the Korea study as being run independently in Korea under very similar conditions as the pivotal study (study population, treatment protocol, follow-up period, etc.); the only major difference noted being that the Korea study was conducted in Korean while the Pivotal Study was conducted in English or Hebrew. The Korea Study was a single-center, randomized (2:1, treatment:sham), double blind, sham controlled trial using the neuroAD.

The Korea study did not provide details regarding the analysis populations (e.g., primary effectiveness (PE) or Per Protocol (PP)) but 26 of the 27 subjects were included in the efficacy results. Results of the Korea Pilot study alone at 7 weeks and 12 weeks are presented below in separate figures for the entire cohort and the neuroAD-defined subgroup of baseline ADAS-Cog≤30. At both timepoints and for both cohorts (entire cohort and the subgroup restricted to baselined ADAs-Cog≤30 post-hoc) the results favor the active group by margins of almost three points. However, contrary to the pivotal study, the overall cohort performs better than the subgroup and the maximum differences are noted at 7 weeks and not 12 weeks.

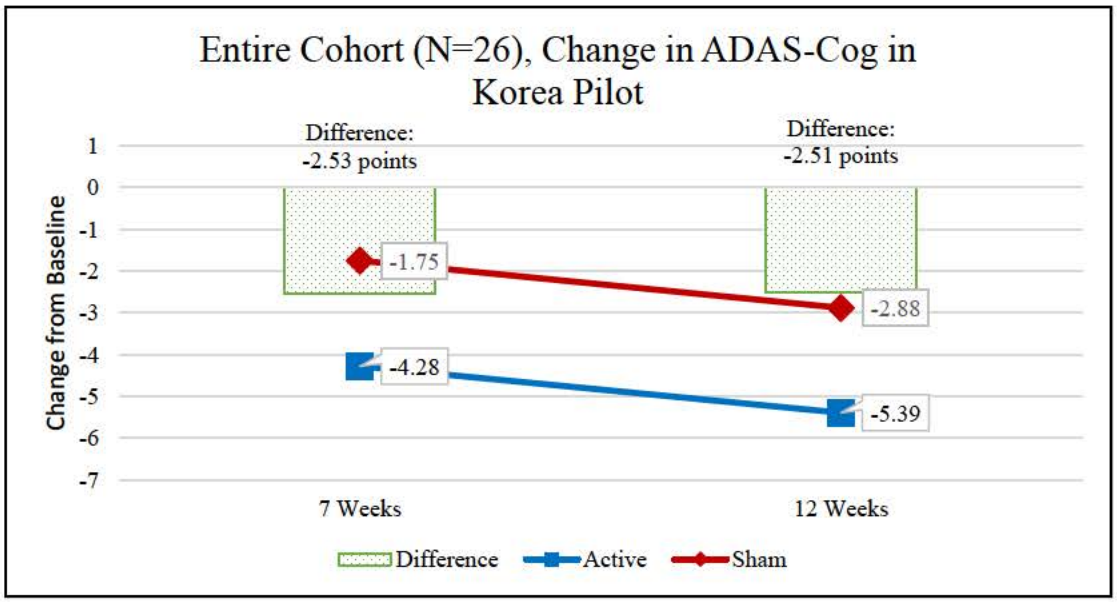


Figure 46. Korea Pilot - Change in ADAS-Cog, Entire Cohort, 7 and 12 weeks

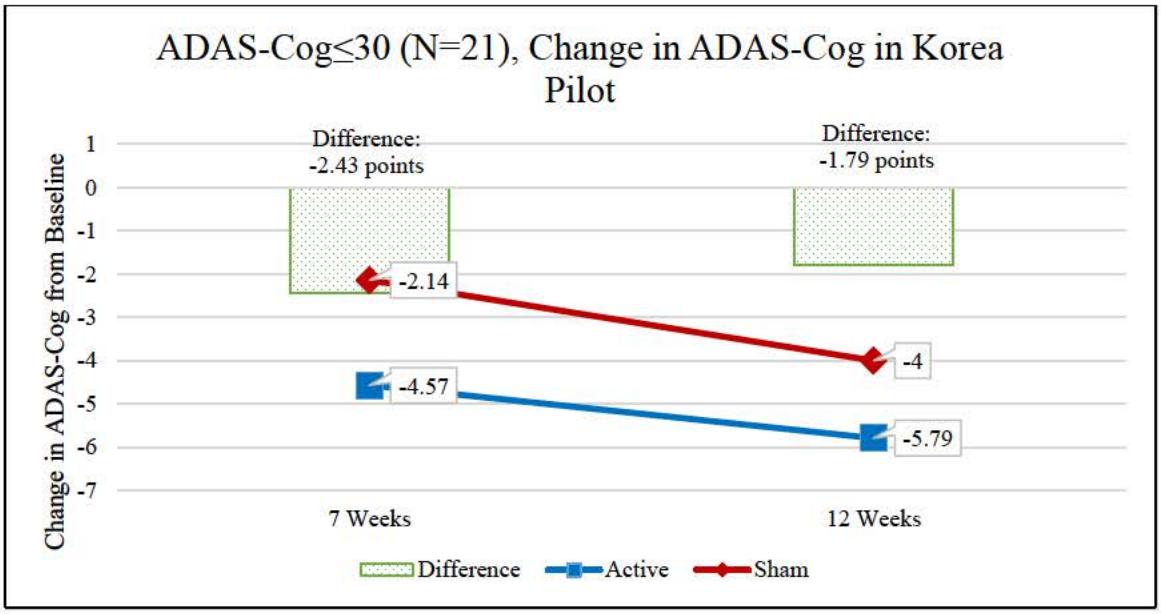


Figure 47. Korea Pilot - Change in ADAS-Cog, Subgroup, 7 and 12 weeks

Pivotal Study Combined with Korea Study at 7 and 12 weeks

In the figures that follow, the results of the pivotal study, the Korea Pilot study, and the pooled results of the two studies are shown; there are separate figures for week 7 results and week 12 results. As shown above, at both timepoints the Korea study shows much better results in favor of the neuroAD treatment. The pooled result is therefore improved from the pivotal study as well.

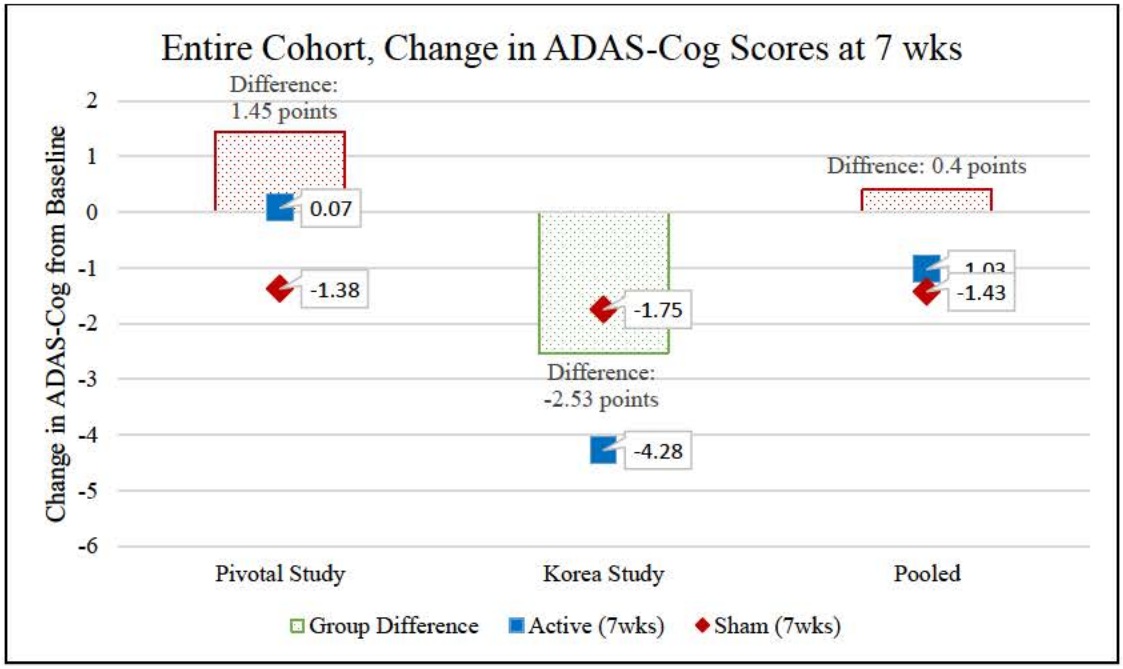


Figure 48. Pivotal Study Post-Hoc Analysis - Korea Pooled Results, Entire Cohort, 7wks

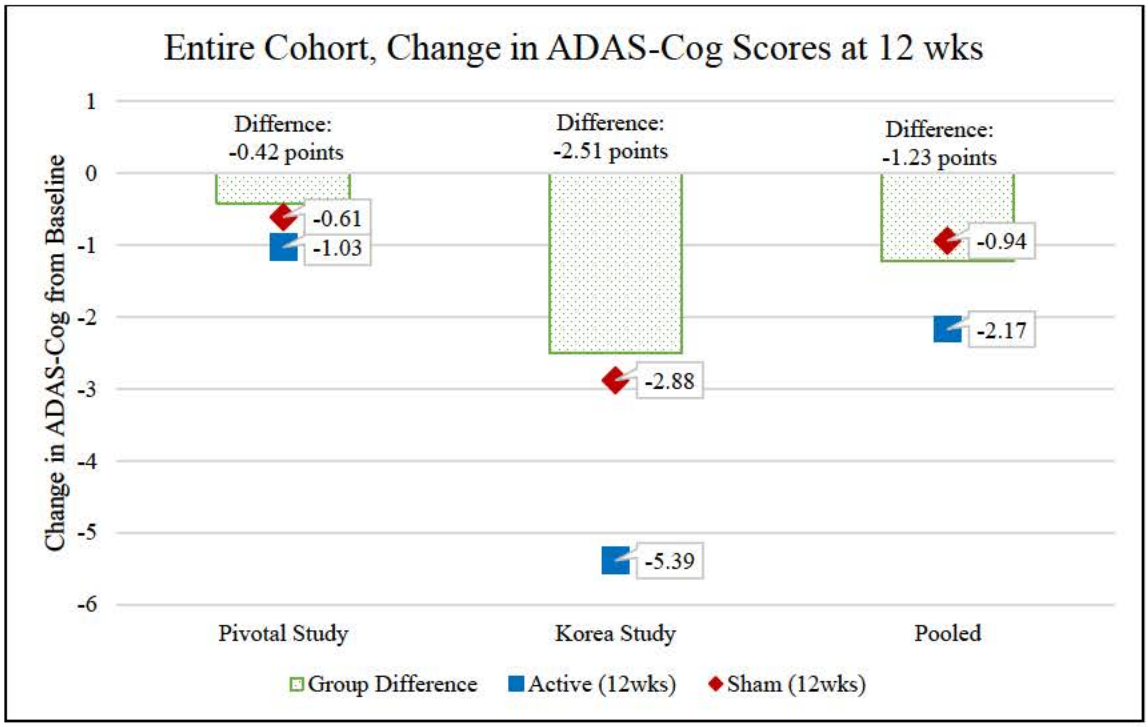


Figure 49. Pivotal Study Post-Hoc Analysis - Korea Pooled Results, Entire Cohort, 12 wks

Additionally, given the results observed related to baseline ADAS-Cog score in the post-hoc analysis of the Pivotal Study data, the combined data were further stratified by baseline score of ADAS-Cog score ≤30.

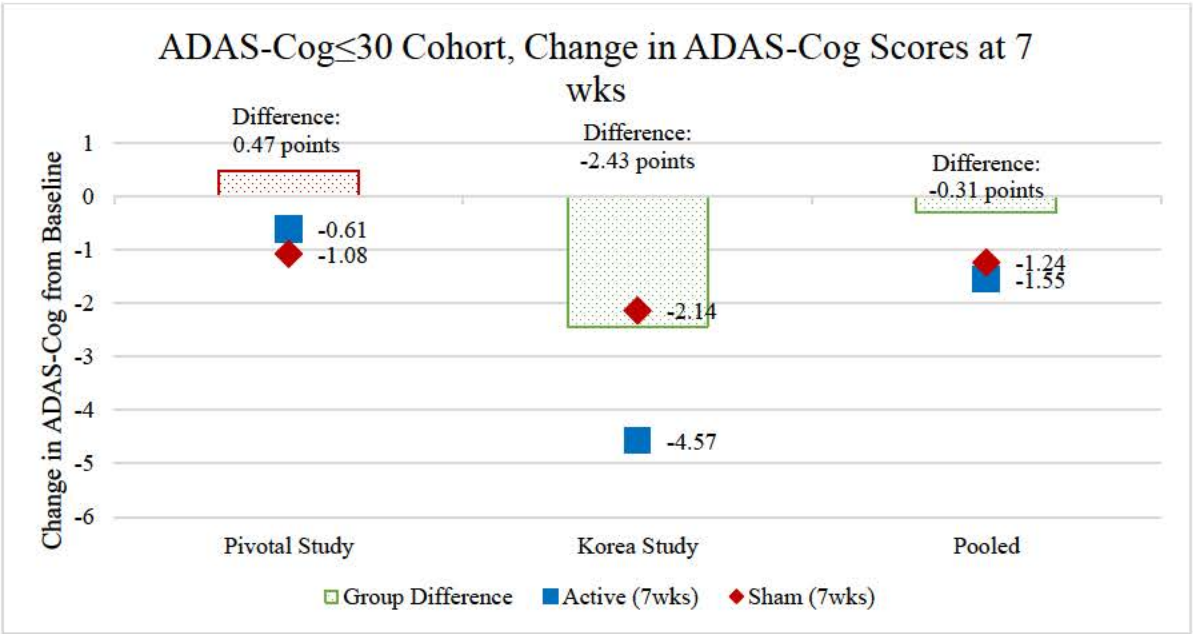


Figure 50. Pivotal Study Post-Hoc Analysis - Korea Pooled Results, Subgroup, 7 wks

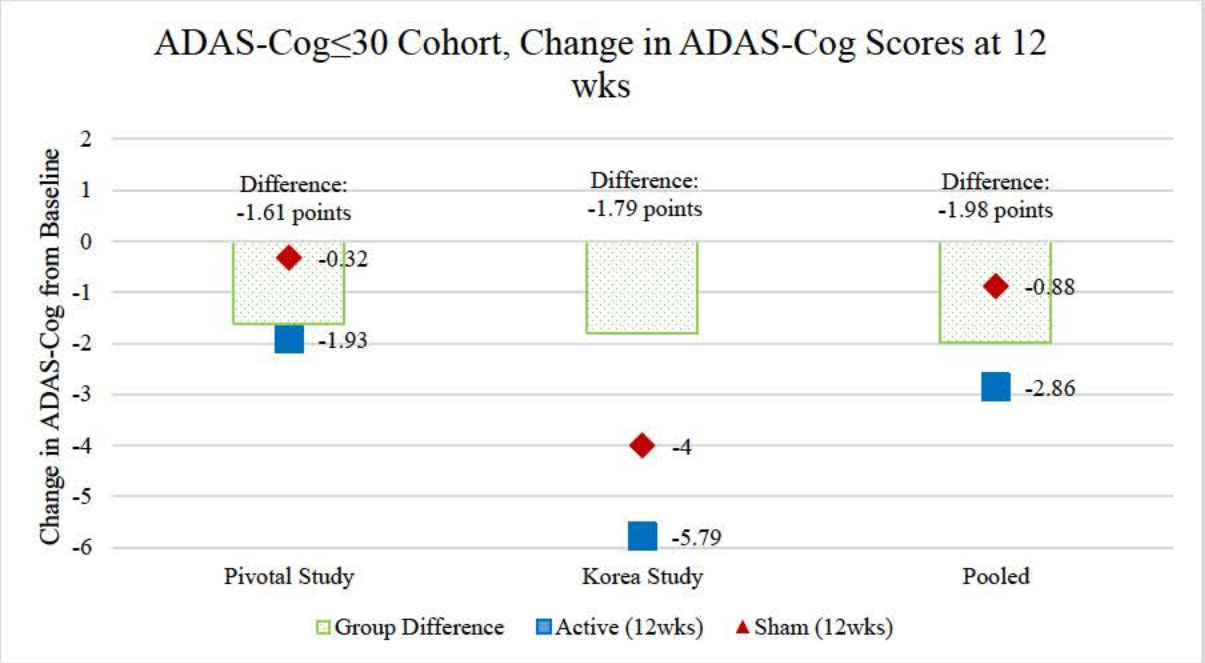
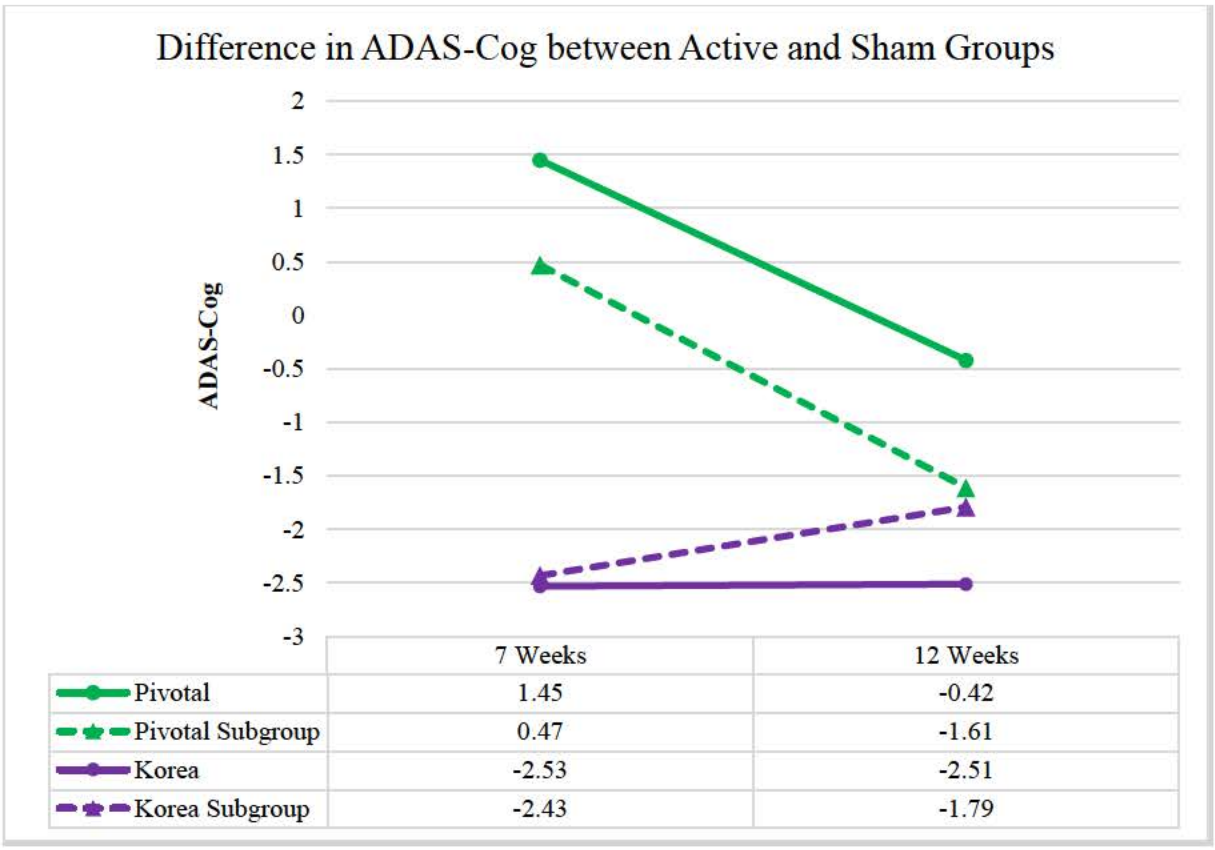


Figure 51. Pivotal Study Post-Hoc Analysis - Korea Pooled Results, Subgroup, 12 wks

Korea Pilot as a Validation of Post-Hoc Analysis of Pivotal Study

As discussed in the section above, it was considered that the Korea Pilot study may represent an independent study that could be used to validate the post-hoc findings of the U.S. pivotal study. However, unlike the neuroAD pivotal study, in the Korea Pilot the between group differences at 7 weeks and 12 weeks for both the entire cohort and the post-hoc subgroup of baseline ADAS-Cog≤30 favored treatment (and the subgroup actually performed worse on average than the entire cohort). Further contradicting the U.S. pivotal study, in the Korea Pilot the trend of the difference between groups does not suggest improvement from Week 7 to Week 12 but instead shows maintenance in the entire cohort and slight worsening in the ADAS-Cog≤30 cohort. These trends are shown graphically in the figure below. It is not clear why similar trends are not shown in these similar study designs using the same device, treatment paradigm, and patient population. This raises concerns regarding the validity of the pivotal study post-hoc analysis and the poolability of these studies.



	7 Weeks	12 Weeks
Pivotal	1.45	-0.42
Pivotal Subgroup	0.47	-1.61
Korea	-2.53	-2.51
Korea Subgroup	-2.43	-1.79

Figure 52. Pivotal Study Post-Hoc Analysis – Using Korea to Validate Pivotal Subgroup

Appendix IV. Korea Pivotal Study and Meta-Analysis (Korea Pilot+Korea Pivotal+US Pivotal)

In Supplement 2 of the de novo review the sponsor provided interim results from 22 subjects (11 active, 11 sham) enrolled in the Korea Pivotal Study (Korea-2). The Korea Pivotal study was designed to limit entry to only those with baseline ADAS-Cog 17-30 based on the investigator conclusions in the Korea Pilot study that "[t]he present results suggest that rTMS-COG represents a useful adjuvant therapy with cholinesterase inhibitors, particularly during the mild stage of AD." (Lee et al., 2016)

In the Korea Pivotal study, the Active group reported a mean change in ADAS-Cog of -3.09 points at 7 weeks and -3.64 points at 12 weeks. The Sham group reported a mean change in ADAS-Cog of -0.55 points at 7 weeks and -1.91 points at 12 weeks. The between-group difference for mean change in ADAS-Cog was statistically significant at the 7-week follow-up visit (p=0.03. Though the Active group reported improvement between the first and second follow-up visit, the between-group difference at 12 weeks was -1.7 points which did not maintain statistical significance. Again, this decrease in magnitude of change between groups over time contrasts that seen in the US Pivotal Study.

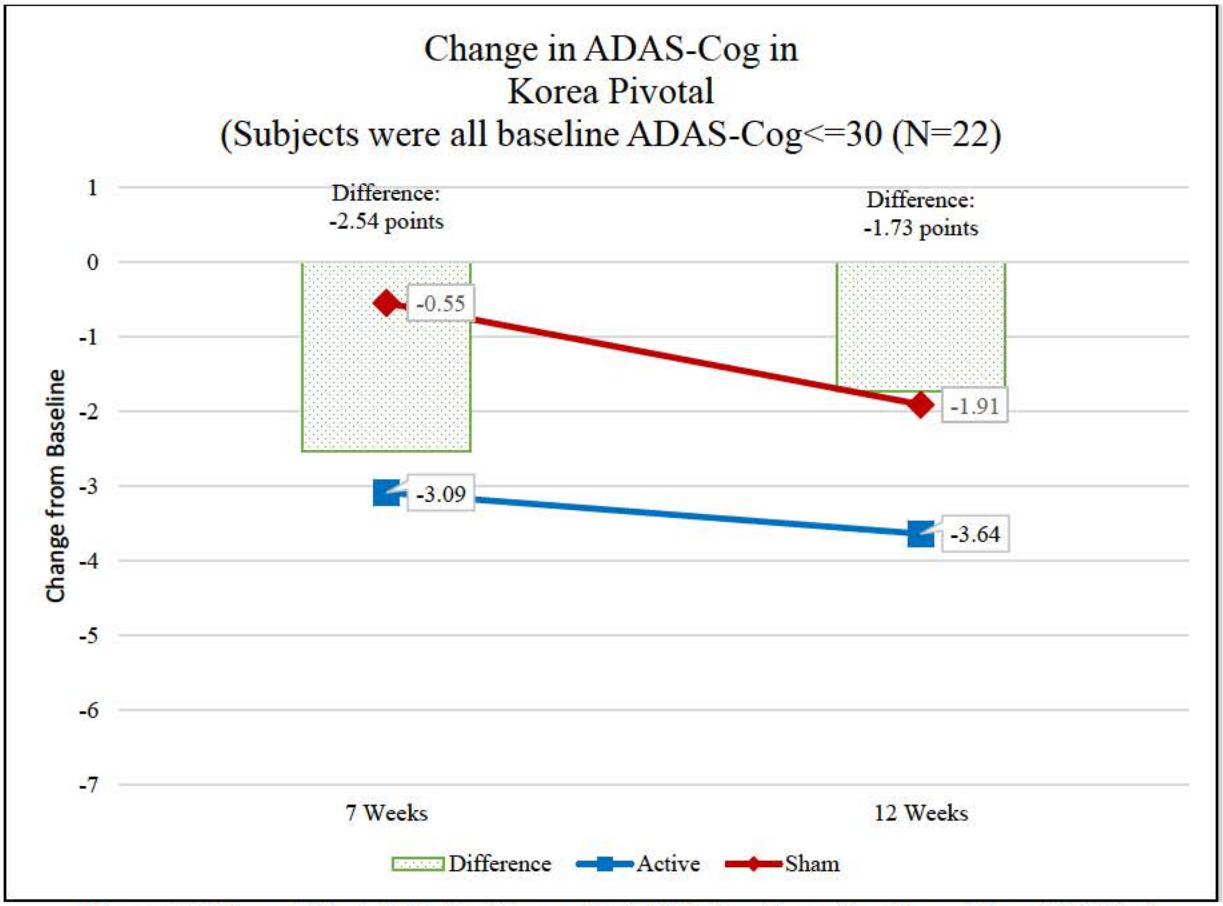


Figure 53. Korea Pivotal Study, Change in ADAS-Cog from Baseline at 7 and 12 Weeks

In Section 6.3.4 of the sponsor Executive Summary, the sponsor provides a meta-analysis which combines the results from the Korea Pilot Study, the Korea Pivotal Study, and the US Pivotal study. The data from the Korea Pilot and the US Pivotal were limited to the baseline ADAS-Cog<=30 subgroup while the Korea Pivotal study only enrolled subject that met this criterion and was included in its entirety. We do not display this analysis here and instead caution the panel when interpreting the results of this analysis (please also see FDA Summary Comments on Pivotal Study Post-Hoc Analysis ADAS-Cog≤30 Subgroup where FDA performs analyses using these datasets).

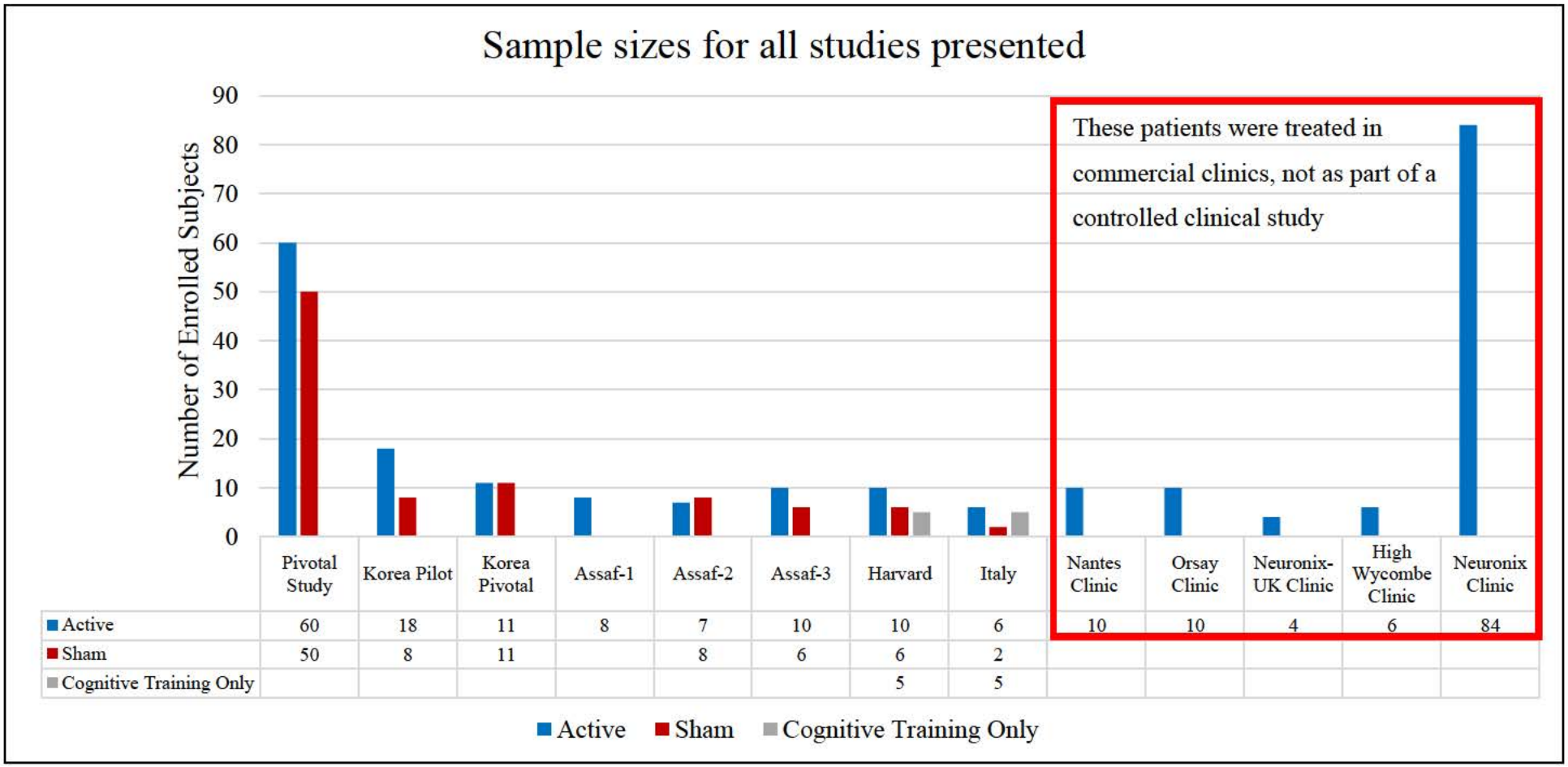
Because the protocols were similar among the three studies, the sponsor assumed that the studies could be considered exchangeable, and therefore included together within a meta-analysis. The assumption of exchangeability may be questioned due to subjects being enrolled in different countries with different languages and cultures. Tests for poolability and assessments for exchangeability of data between OUS and US sites is typically prespecified in the Statistical Analysis Plan. Though FDA has not specifically reviewed the data that is presented in the meta-analysis for exchangeability, we note that according to the sponsor there were differences in baseline motor thresholds (MT) across the studies. In the Korean pilot study, for example, the average MT at baseline for all active patients was 94.4%, which is significantly higher than the average MT at baseline for all active patients from the neuroAD pivotal study (74.2%). Also, for the 4 active patients in the Korean pilot study who had baseline ADAS-Cog > 30, the average motor threshold was 93%. Because the intensity of TMS is typically adjusted according to the patient's baseline MT, the intensity of TMS treatment may differ between the Korea pilot study and the US pivotal study. This is one such example of uncertainty with regards to this meta-analysis.

Appendix V. Post Hoc Analyses Using Supplemental Clinical Data Sources

Neuronix provided Supplemental Clinical data sources in response to Deficiency #1 of the 3/13/2017 AINN letter. Some of these supplemental data sources were also listed as "pilot studies" in the original submission. In the de novo supplement, Neuronix pooled these supplemental clinical data sources to provide additional clinical data in support of the neuroAD effectiveness. **The results of the pooled data do not include the Pivotal Study**.

Note: The analyses presented in this section are post-hoc analyses without pre-specification and multiplicity adjustment. Therefore, we do not include p-values.

The supplemental data sets included patients treated at commercial clinicals and a number of small studies. All but one of the datasets (Harvard) was derived from subjects residing outside the United States (OUS). For comparison, the sample sizes for all the clinical data sets are shown in the figure below.



	Pivotal Study	Korea Pilot	Korea Pivotal	Assaf-1	Assaf-2	Assaf-3	Harvard	Italy	Nantes Clinic	Orsay Clinic	Neuronix-UK Clinic	High Wycombe Clinic	Neuronix Clinic
Active	60	18	11	8	7	10	10	6	10	10	4	6	84
Sham	50	8	11		8	6	6	2					
Cognitive Training Only							5	5					

Figure 54. Supplemental Data Sets - Sample Size Comparison

Results of Pooled Supplemental Investigations

These results include pooled data separated into the overall cohort and the ADAS-Cog≤30 subgroup. Results include new ADAS-Cog analyses pooled from all available data at a 6-10 week post-intervention timeframe and a 10-14 week post-intervention timeframe. Neuronix also provided new CGI-C results available from two of the studies that had already been reported to FDA (Assaf-3 and Korea Pilot) for the 10-14 week timeframe.

Pooled ADAS-Cog Results

A summary of all the data with respect to ADAS-Cog is provided below. At the first follow-up, data from 180 subjects are available (139 active subjects and 41 sham subjects). At the second follow-up period, data from 65 active subjects and 27 sham subjects is available.

Table 22. Supplemental Investigations Contributing to Pooled ADAS-Cog Analysis

Study	Subjects with Change in ADAS-Cog Data at 6-10 Weeks		Subjects with Change in ADAS-Cog Data at 10-14 Weeks	
	Active	Sham	Active	Sham
Assaf-1	8	0	0	0
Assaf-2	7	8	0	0
Assaf-3	10	6	10	6
Harvard (Beth Israel)	10	6	10	0
Korea Pilot	18	8	18	8
Korea Pivotal	11	11	11	11
Italian	6	2	6	2
Nantes Clinic	10	0	0	0
NeuroCare Clinic Israel	59	0	10	0
Total	**139**	**41**	**65**	**27**

At the 6-10 week follow-up period the neuroAD™ group reported mean change in ADAS-Cog score of -2.26 while the sham group reported a mean change of -0.18 points. The between-group difference reported at the 6-10 week follow-up window is -2.08 points in favor of the active group.

At the 10-14 week follow-up window, the neuroAD™ group reported a mean change in ADAS-Cog score of -3.56 points while the sham group reported a mean change of -1.4 points. The between-group difference reported at the 10-14 week follow-up was -2.16 in favor of the active group.

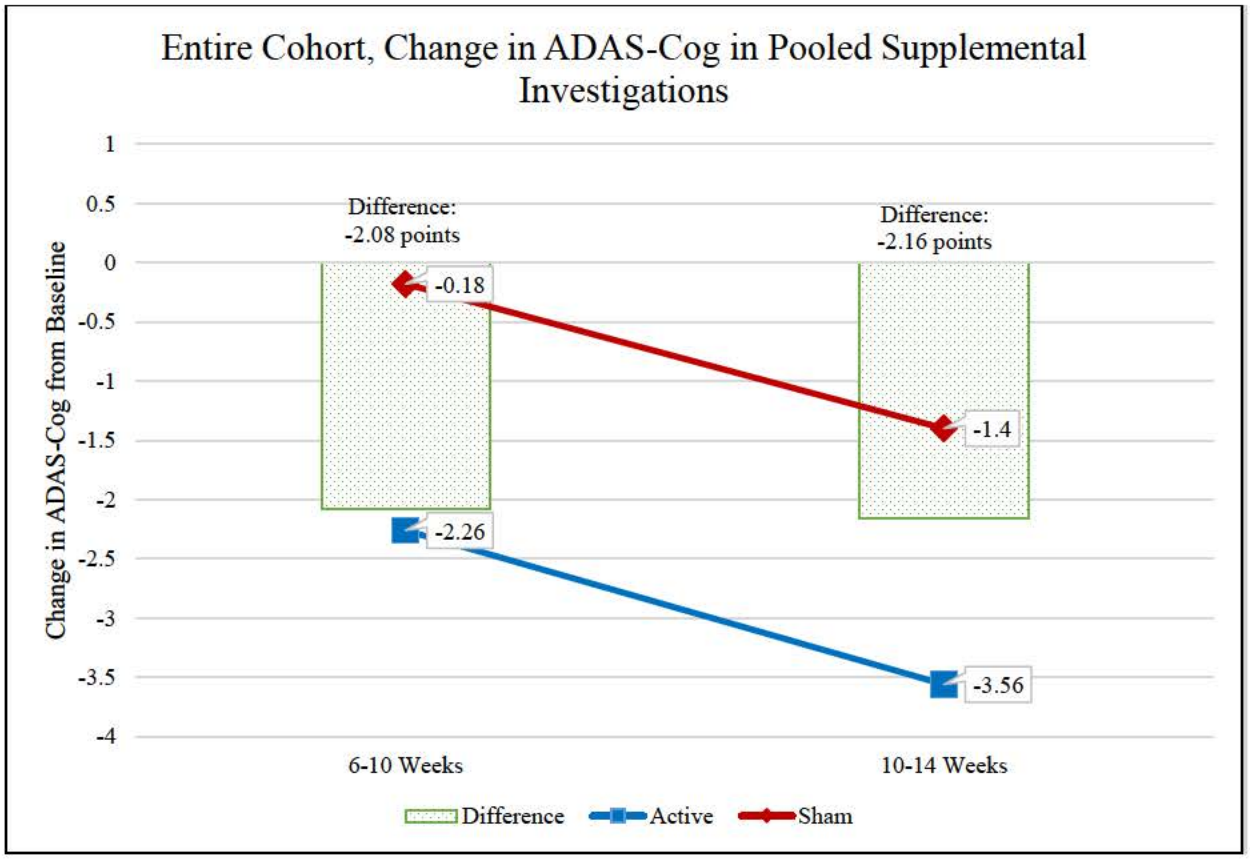


Figure 55. Supplemental Investigations, Pooled ADAS-Cog Group Changes from Baseline

New ADAS-Cog analyses were also provided for the ADAS-Cog≤30 subgroup that was proposed from the post-hoc analysis of the pivotal study data. For the subgroup, at the first follow-up Neuronix states that there is data from 157 subjects (121 active subjects and 36 sham subjects). At the second follow-up period, data from 79 subjects are available (55 active subjects and 24 sham subjects).

At the 6-10 week follow-up period the neuroAD group reported mean change in ADAS-Cog score of -2.07 while the sham group reported a mean change of -0.03 points. The between-group difference reported at the 6-10 week follow-up window is -2.07 points in favor of the active group.

At the 10-14 week follow-up window, the neuroAD™ group reported a mean change in ADAS-Cog score of -3.37 points while the sham group reported a mean change of -1.61 points. The between-group difference reported at the 10-14 week follow-up was -1.76 in favor of the active group.

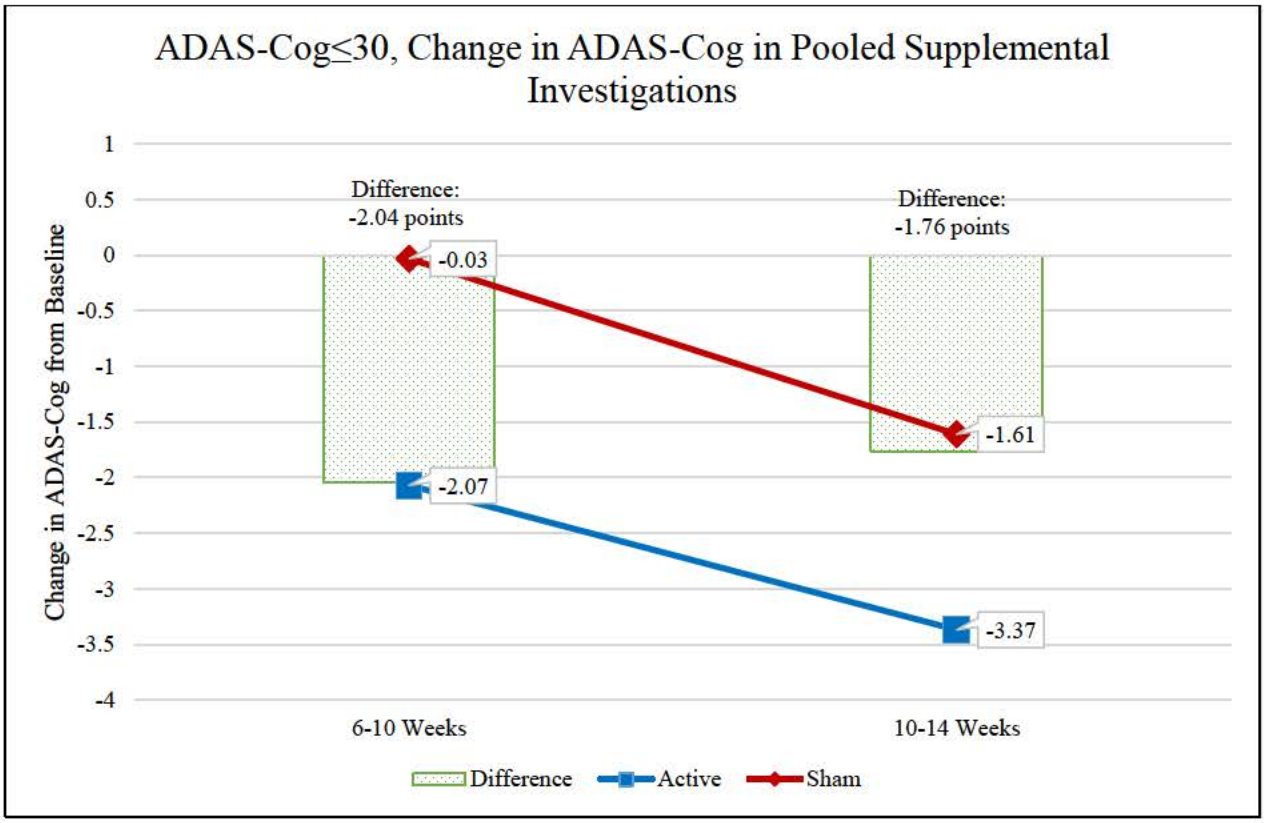


Figure 56. Supplemental Investigations, Subgroup, Pooled ADAS-Cog Group Changes from Baseline

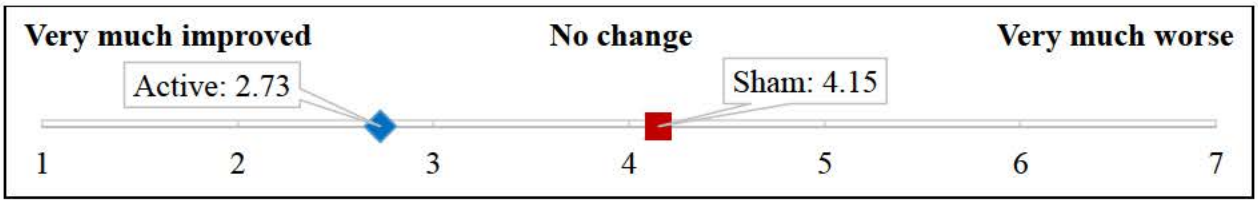
Pooled CGI-C Results

Change in CGI-C from baseline was assessed in the combined sample at the 10-14 week time point for all studies where data was available. The two studies that reported CGI-C during the applicable period, shown in the table below, have been combined in order to conduct the following CGI-C assessments. **This data is NOT limited to the ADAS-Cog≤30 cohort.**

Table 23. Supplemental Investigations Contributing to Pooled CGI-C Analysis

Study	n	Subjects with Change in CGI-C Data at 10-14 Weeks	
		Active	Sham
Assaf 3	13	8	5
Korean pilot study	26	18	8
Total	**39**	**26**	**13**

The combined active group reported mean change in CGI-C score of -1.27 points (i.e., improvement) while the combined sham group reported mean change in CGI-C score of 0.15 points (i.e., deterioration). The between group difference for reported change in mean CGI-C was -1.42 in favor of the active group.



FDA Summary Comments on Pooled Supplemental Investigation Results

FDA has conducted independent analyses on the available raw data from the Supplemental Investigations. Table 24 below provides a summary of the available data by percentage of total data points based on the number of subjects.

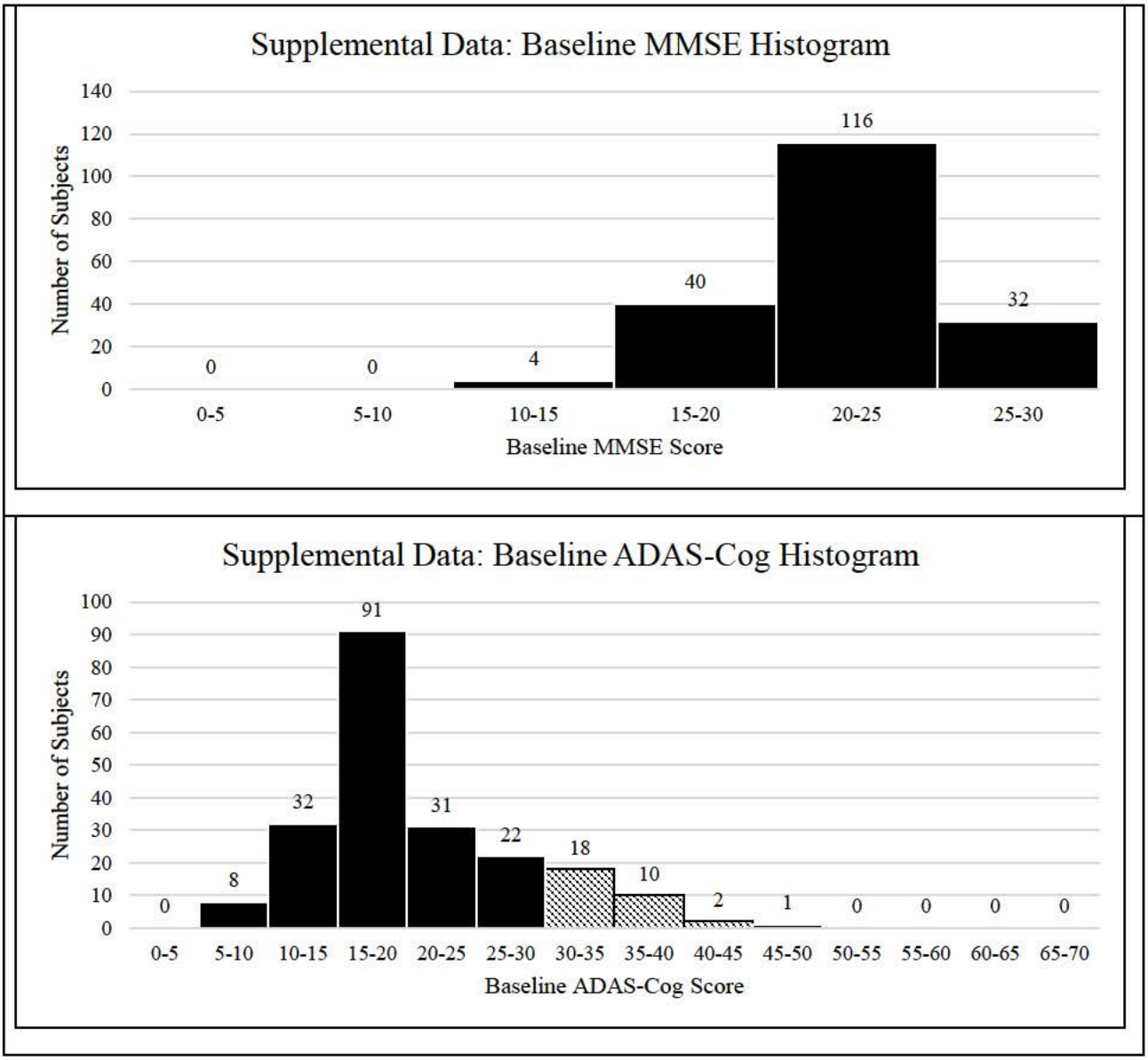
Table 24. Source and Quantity of Raw Data Available from Supporting Investigations

Study Name	N	Reference	Available Raw Data % (data/total entries)					
			Safety++	ADAS-Cog Baseline	ADAS-Cog 6-10wk	ADAS-Cog 10-12wk	CGI-C 6-10wk	CGI-C 10-14wk
Randomized Controlled Trials								
Assaf-2*	15	Published article (Rabey et al., 2013)	0% (0/15)	100% (15/15)	100% (15/15)	86% (13/15)	100% (15/15)	86% (13/15)
Harvard+	21	Unpublished article	100% (21/21)	100% (21/21)	100% (21/21)	71% (15/21)	0% (0/21)	0% (0/21)
Assaf-3	16	Unpublished article	0% (0/16)	100% (16/16)	100% (16/16)	100% (16/16)	93% (15/16)	81% (13/16)
Korea Pilot^	27	Published article (Lee et al., 2016)	100% (25/25)	100% (25/25)	100% (25/25)	100% (25/25)	100% (25/25)	100% (25/25)
Korea Pivotal	22	Source information not provided	100% (22/22)	100% (22/22)	100% (22/22)	100% (22/22)	0% (0/22)	0% (0/22)
Thailand	9	Poster at AAIC-2016 Conference	0% (0/9) for all - No raw data provided					
Italy	13	Source information not provided	100% (13/13)	100% (13/13)	100% (13/13)	100% (13/13)	0% (0/13)	0% (0/13)
Open Label Trials								
Assaf-1	8	Published article (Bentwich et al., 2011)	0% (0/8)	100% (8/8)	100% (8/8)	87% (7/8)	100% (8/8)	12% (1/8)
France (Nantes)	10	Published article (Nguyen et al., 2017)	100% (10/10)	100% (10/10)	100% (10/10)	100% (10/10)	0% (0/10)	0% (0/10)
Commercial Clinics								
Israel** (NeuroCare)	84	Unpublished article	100% (84/84)	100% (84/84)	77% (65/84)	21% (18/84)	0% (0/84)	0% (0/84)
France*** (Orsay)	10	Source information not provided	0% (0/10)	0% (0/10)	0% (0/10)	0% (0/10)	0% (0/10)	0% (0/10)
UK***	10	Source information not provided	0% (0/10)	0% (0/10)	0% (0/10)	0% (0/10)	0% (0/10)	0% (0/10)

*1/15 subjects changed medication during the study. Subject included in analysis and has data for all assessment scales and timepoints
**5/84 subjects in NeuroCare (Israel) dataset were undergoing their second treatment. 1/5 of those subjects also participated in an Assaf study
***The Orsay and UK clinics did not administer the ADAS-Cog and did not provide safety data
^Only 25 data entries were provided though study sample size is listed as N=27
+Safety data was reported from the site in aggregate, not by individual subject
++ **This table shows the availability of raw data for the Supplemental Investigations that was provided to FDA on November 16th, 2018. As such, a 0% does not indicate no AEs were recorded but indicates that no raw data was provided.**

In general, FDA finds the pooled result from the supplemental investigations to carry significant uncertainty due to the quality of the evidence presented. The lack of safety data provided for the Assaf studies also contribute to incomplete safety data when making a benefit/risk assessment.

Overall, the FDA defined and extracted the following relevant baseline demographics of the Supplemental Investigations from the raw data provided by Neuronix.



Though there is significant uncertainty in the pooled efficacy results, FDA finds it beneficial to use these Supplemental Investigations to provide validation of the post-hoc analysis of the pivotal study (the most robust dataset).

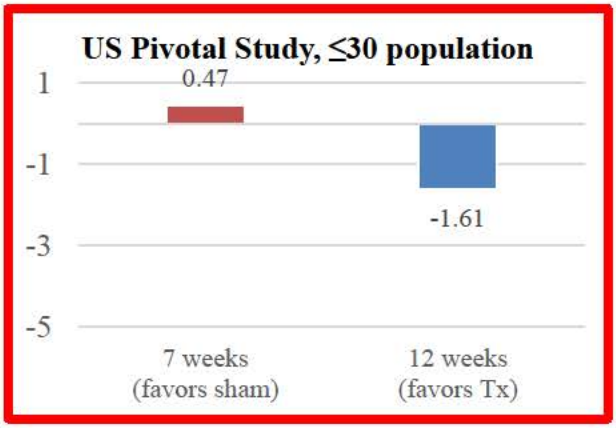
Investigating Post-Hoc Assessment Finding at 12-Weeks

First, we look to the supplemental investigations to compare the trends in the treatment effect over time. It was an unexpected result from the pivotal study that the results in favor of the device were only noted at the 12 week endpoint (6 weeks after the treatment ended). The plots below show the between-group differences in the ADAS-Cog improvement over time.

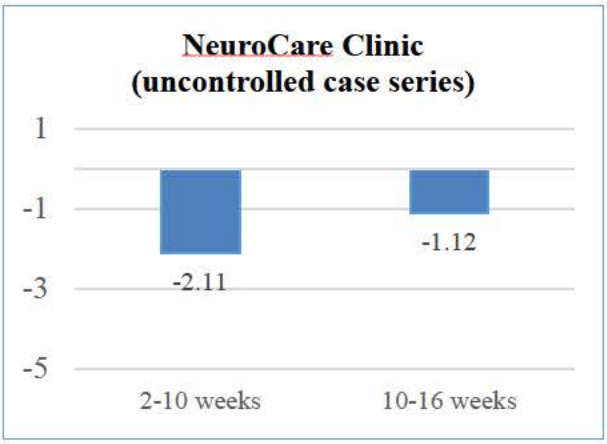
Table 25. Supplemental Investigations - Between-group differences in ADAS-Cog over time

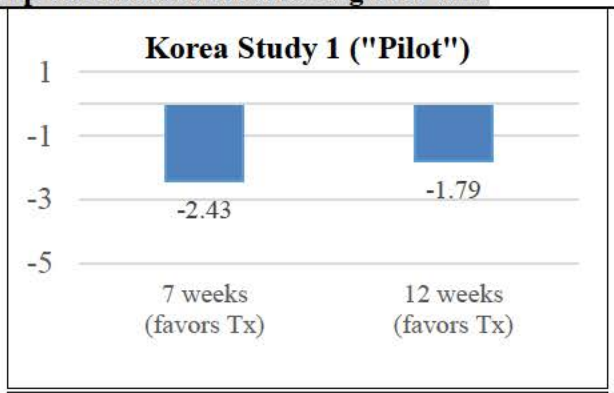
The two data points in each of these charts represents the difference in the ADAS-Cog score between the active and sham groups recorded at each of the two follow-up visits.

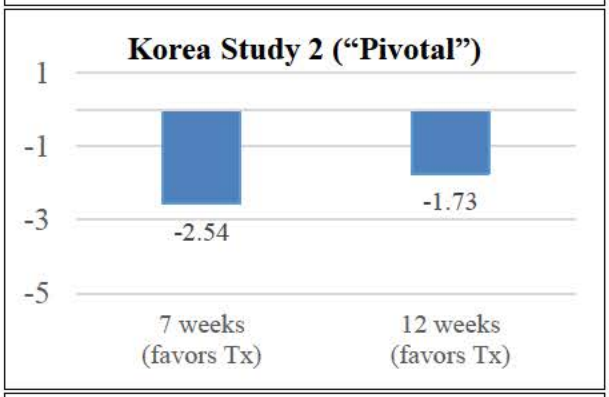
Among the sham-controlled studies where there are two recorded follow-up visits and ADAS-Cog data from both groups, the pivotal study is the only dataset where there the difference between the active and sham groups increases over time.

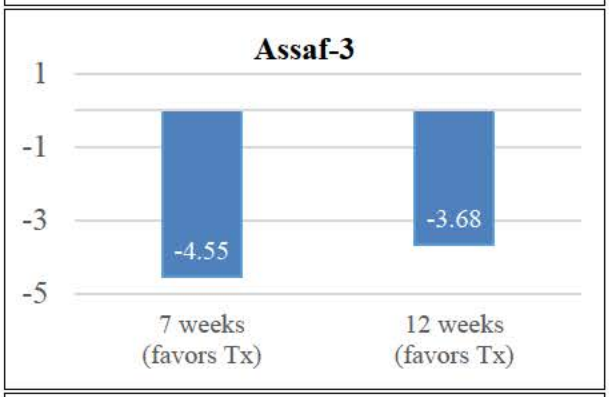


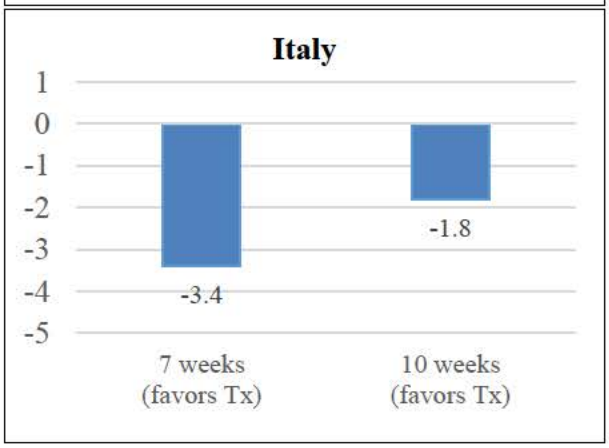
In the other four studies that have data from two follow-up visits to compare over time, the difference decreases. The NeuroCare Clinic dataset, while not sham-controlled, demonstrates a similar trend:





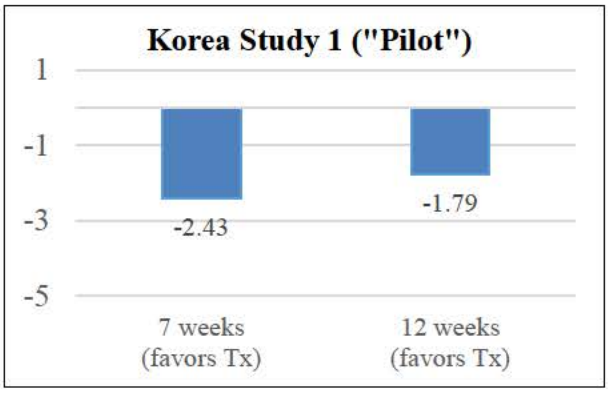
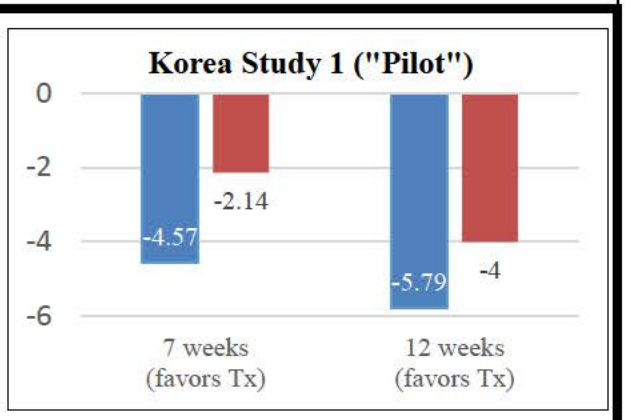
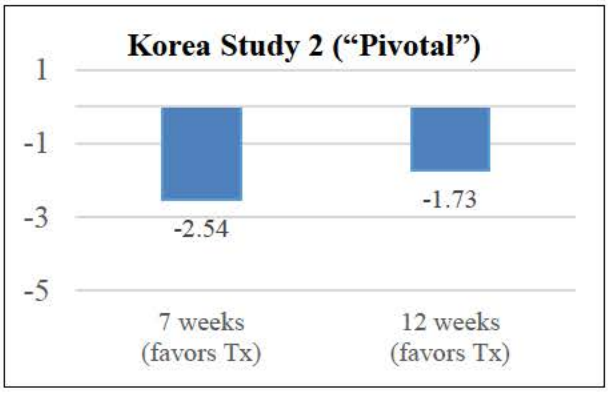
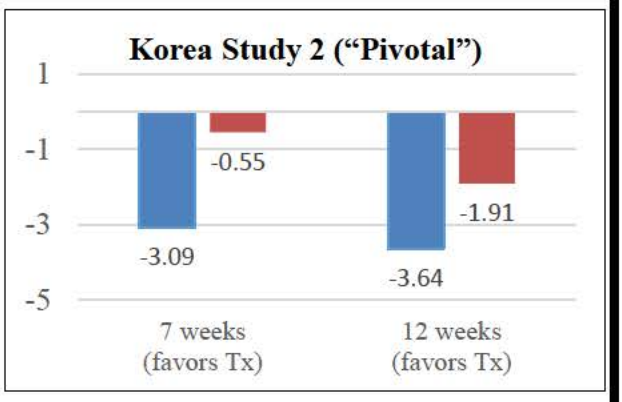
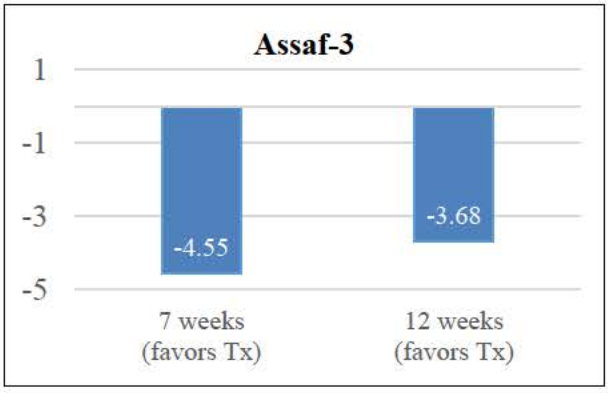
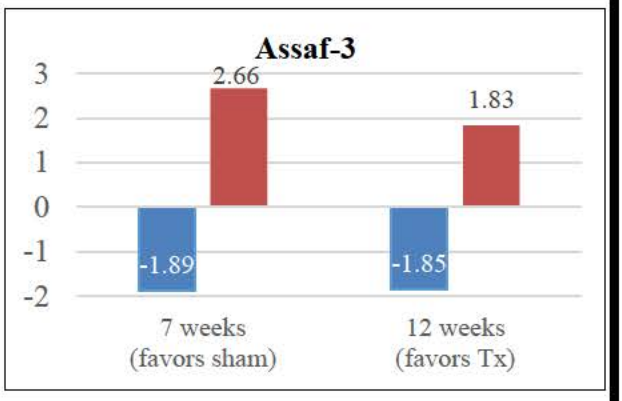


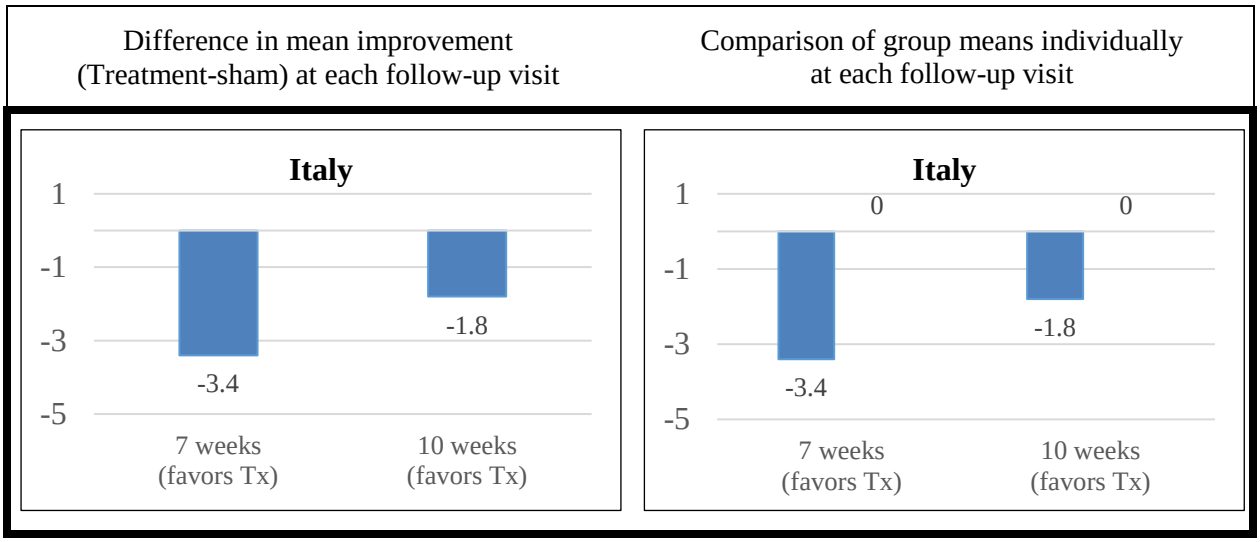




In the plots below we also continue to investigate the differences in the trend between 7 and 12 weeks and note that the active and sham contributions that form these trends are also inconsistent.

Table 26. Supplemental Investigations – Between-group differences in ADAS-Cog over time, comparing Active and Sham

Difference in mean improvement (Treatment-sham) at each follow-up visit	Comparison of group means individually at each follow-up visit
	
	
	



Investigating Post-Hoc Subgroup of best effect seen in those with baseline ADAS-Cog≤30

We then looked to those studies that had data at the 10-14 timepoint in the ADAS-Cog≤30 population to provide comparison/validation of the best-case result in favor of treatment found in the pivotal study post-hoc analysis (difference between groups was -1.61). Neuronix provided the summary results for the Pooled Supplemental dataset that conforms to these requirements that we have used in the figure below. To the best of our knowledge the studies that conformed to the data requirements and therefore were used in the analysis were Korea Pilot, Korea Pivotal, Assaf-3, and Italy. Comparative results with the pivotal post-hoc analysis are shown in the figure below.

Of interest, it is noted that the best-case result in favor of the active group from the pivotal post-hoc analysis (-1.61) was identical to the sham result in the same subgroup at the same time range in the pooled investigations.

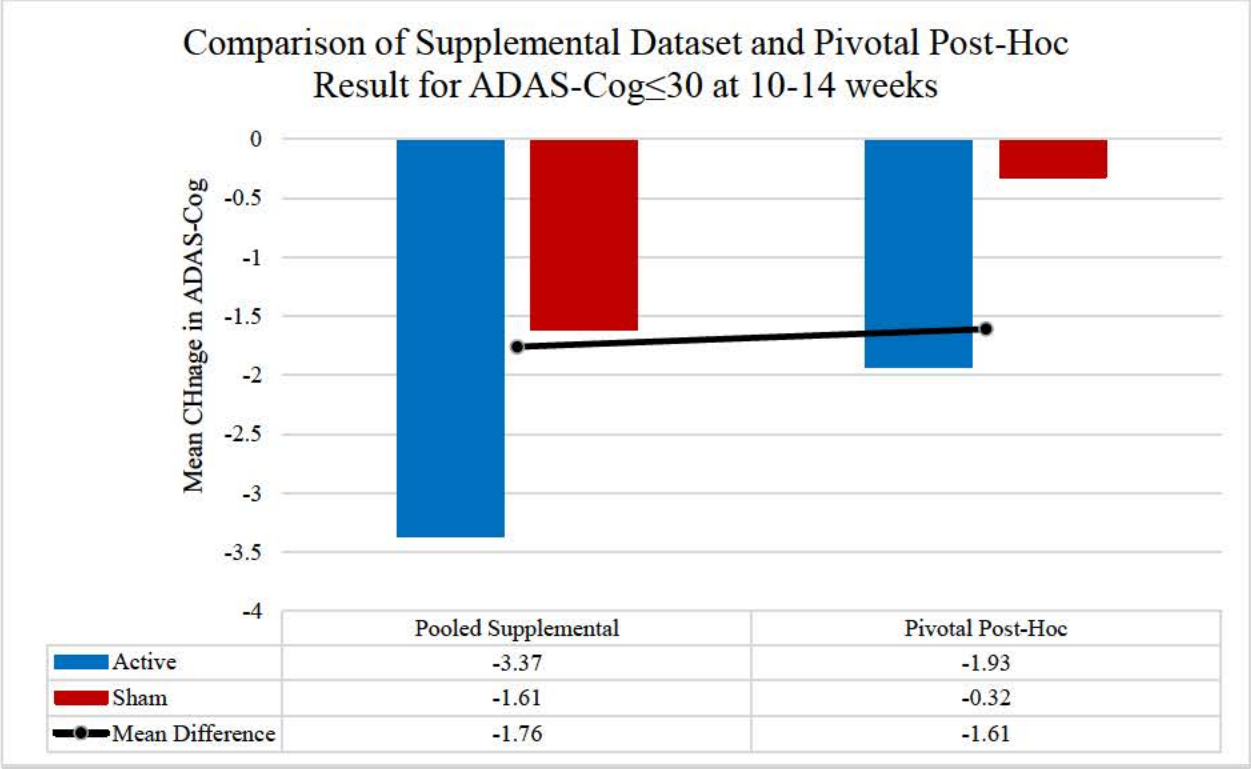


Figure 57. Comparison of ADAS-Cog Pooled Supplemental Results Limited to Subgroup and Pivotal Post-Hoc Subgroup at 10-14wks

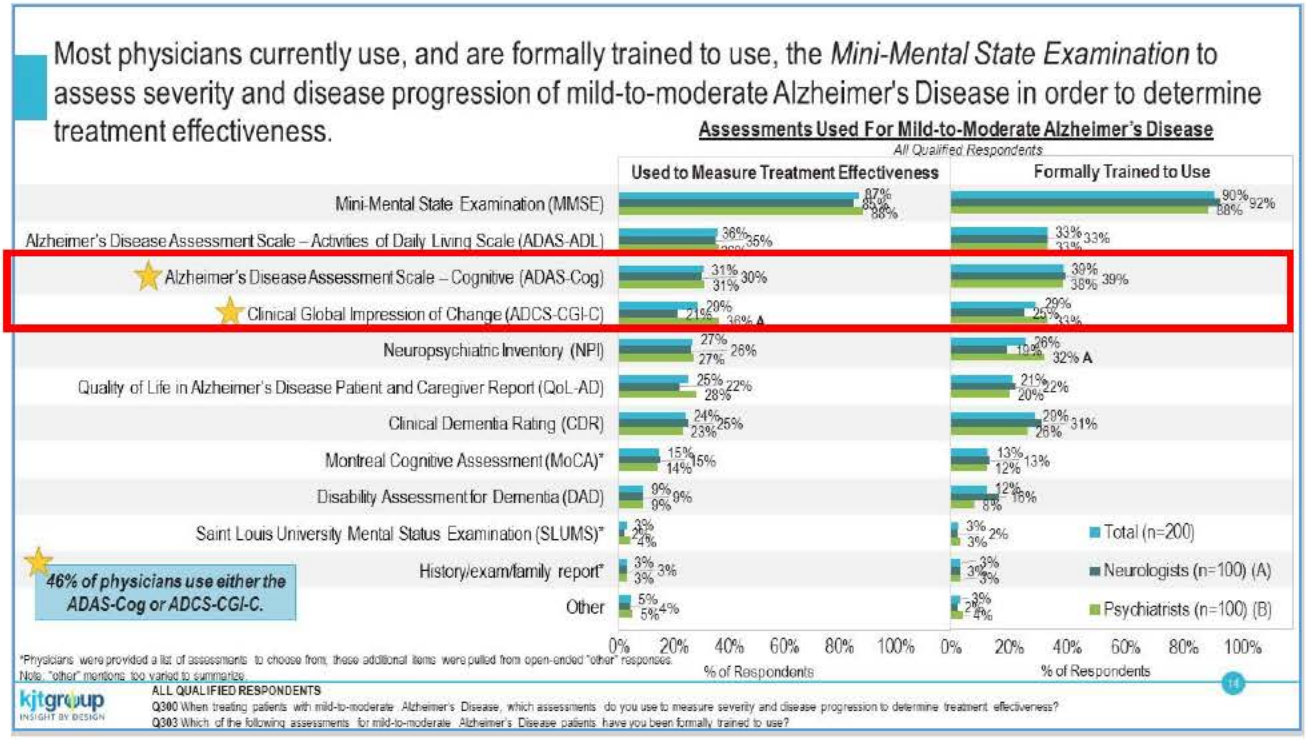
Looking at just the mean difference between groups, the results of the pooled investigations do appear to provide some validation to the post-hoc analysis of the pivotal study as the mean differences are fairly consistent between the two datasets -1.76 and -1.61. However, in Table 26 above, we have demonstrated that the active and sham group results that contribute to these mean group differences for each of the individual studies behave very differently. Specifically, the active and sham groups show different result trends in the two Korea studies than they do in the Assaf-3 study.

Appendix VI. Stakeholder Input

Physician Survey

Neuronix provided FDA with the methodology and results of a physician survey conducted by an independent party (KJT Group) between February 23 and March 16, 2018. This survey was performed without any request from FDA though it was designed to address the FDA concern regarding clinical meaningfulness of the ADAS-Cog results. Per Neuronix, the intent of the survey was to obtain feedback from neurologists and psychiatrics regarding what they would consider a clinically meaningful change on the ADAS-Cog scale after an intervention. Neuronix concluded the following: *"Nearly half of physicians consider at least a 1 point improvement (or less, so long as there is no deterioration) in ADAS-Cog Score clinically meaningful following 3 months of treatment; even more find this threshold clinically meaningful when there is also a 0.5 point improvement in ADCS-CGI-C."*

However, FDA is concerned about the qualifications/training of the physicians who answered the question regarding a *clinically* meaningful change on the ADAS-Cog scale. Based on summary information presented by Neuronix, only 31% of Respondents use the ADAS-Cog for which they are providing an assessment of *clinical* meaningfulness. Likewise, only 29% use the CGI-C which is used in a follow-up question as evidence to support a smaller change in ADAS-Cog as a potential MCID..

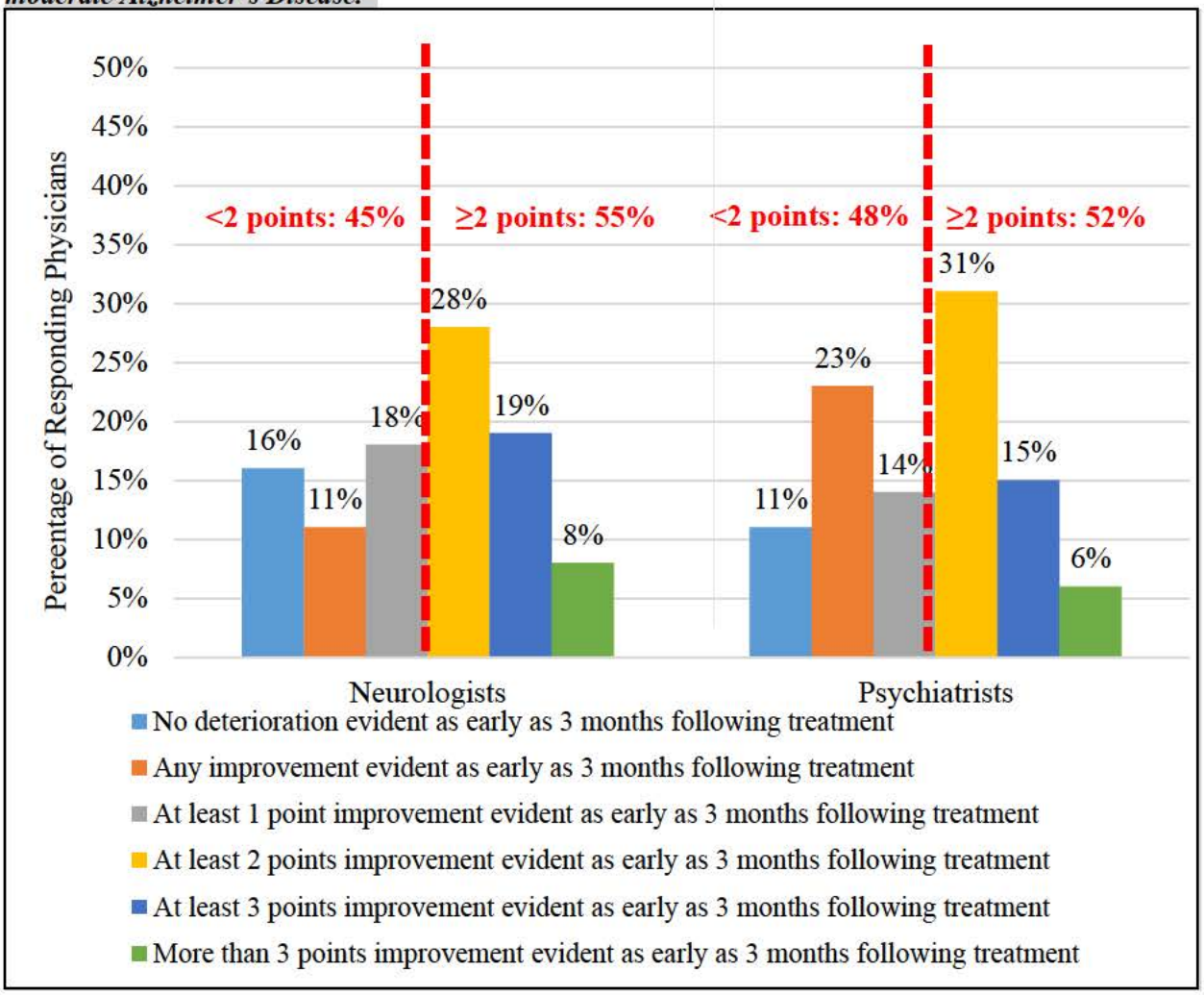


These concerns aside, Neuronix reports the results of the survey as are as follows:

- Without an additional CGI-C improvement, about one third of physicians (30%) responded that at least a 2-point improvement on the ADAS-Cog is clinically meaningful. This 30% of responders at the 2-point cut-off value was the largest percent of consensus among the group.

FDA created the chart below to visualize the survey results provided to FDA by Neuronix.

Survey Question Q310: "Now, again thinking from your perspective, what minimal degree of improvement in the ADAS-Cog score would you consider clinically meaningful when evaluating the effect of a new treatment administered on top of Cholinesterase inhibitors to your patients with mild-to-moderate Alzheimer's Disease."



Based on the Physician Survey results, FDA concludes that even when considering an adjunct therapy more than half of physicians considered at least 2 points or greater on the ADAS-Cog score to be clinically meaningful following 3 months of treatment.

It may also be considered that the data should be reanalyzed to show the following:

- The percentage of those responding with numerical point improvements only (i.e., removing those who responded "no deterioration" and "any improvement"). "No deterioration" is a response that does not apply to the intervention-based intended use and clinical study designs of the neuroAD. "Any improvement" is not able to be quantified for scientific analysis.

- Responses limited only to those physicians who responded that are trained to use the assessment scales and that they actually *use* the ADAS-Cog and the CGI-C. This is intended to target the population of physicians that are most familiar with rating clinical signs and symptoms using the

ADAS-Cog scale and are therefore in the best position to comment on the clinically important difference of the scale.

Please also note, while the blinded product profile was shown to survey participants after the questions were answers and therefore should not impact the data that has been provided above, FDA has concerns that potential benefits of the device were overstated and the potential risks were understated in the blinded product profile. For example, in the product profile the sponsor stated that the device provided cognitive and functional improvement, and the results presented are those of the per-protocol (PP) population instead of the primary efficacy (PE) population; the PP populations excluded those with major protocol deviations of which 7/8 were poor-performing active subjects. The product profile also did not disclose the primary endpoint of the pivotal study in which the results favored the sham group.

neuroAD (Blinded) Product Profile

- A new, non-invasive, medical device treatment for Alzheimer's Disease, typically administered in combination with pharmacotherapy.
- Treatment is administered at the clinic for one hour per day, 5 times/week, over 6 weeks.
- It utilizes two modalities concurrently combined:
  1. Neuro-navigated focused Transcranial Magnetic Stimulation (TMS) is used to stimulate targeted areas of the brain responsible for various cognitive functions that have been impaired by Alzheimer's disease.
  2. Tailored Cognitive Training is used to target those same areas of the brain while they are being magnetically stimulated.
- It has minimal-to-no side effects, and provides cognitive and functional improvement.
- Results from a pivotal study of patients with a baseline ADAS-Cog score of 17-30 (mild-to-moderate Alzheimer's Disease):

ADAS-Cog Score	Treatment Group subjects experienced a mean improvement in score at 12 weeks of -2.11. The same subpopulation in the placebo group reported a mean change of -0.32 (between groups difference of -1.79 favoring the treatment, statistically significant).
ADCS-CGI-C Score	Treatment Group subjects reported a lower mean CGI-C score than the placebo group with difference between groups of -0.45 (favoring treatment) at week 12 (nearing statistical significance, p=0.07). CGI-C distribution reached statistical significance.

kjtgroup INSIGHT BY DESIGN

Patient/Caregiver Survey

In DEN160053/S002 Neuronix provided results from a SurveyMonkey survey titled "New Treatment for Alzheimer's Disease" which was available between 8/13/2017 and 8/17/2017. This survey was requesting feedback regarding the shortcomings of current treatment and hypothetical new treatments for Alzheimer's disease. The survey was not specific to the evidence available for the neuroAD. Briefly, Neuronix reported 1,073 total participants of which the majority were patient family members and/or caregivers. The question topics were as follows with the bolded question being FDA's focus:

- Sorting questions regarding knowing someone who has Alzheimer's disease
- Caregiver amount of involvement.
- Kinds of treatment given to the Alzheimer patient, side effects and satisfaction from the results.
- Opinion regarding new treatment needed for Alzheimer's disease.
- **Opinions regarding new treatment's conditions with 70% and 50% likelihood of improvement**.
- Demographic information.

The following responses to the bolded were recorded (70% first, then 50%)

"If a new Alzheimer's medical device treatment as available that required 5 visits/week to a clinic for one hour per day, with minimal or no side effects, and a 70% likelihood of some improvement, how likely would you be to want your family member to try it?"

VERY LIKELY	SOMEWHAT LIKELY	NEUTRAL	SOMEWHAT UNLIKELY	VERY UNLIKELY	TOTAL	WEIGHTED AVERAGE
47.31% 44	30.11% 28	13.98% 13	3.23% 3	5.38% 5	93	1.89

"If the same Alzheimer's medical device treatment was available with a 50% likelihood of some improvement, how likely would you be to want your family member to try it?"

VERY LIKELY	SOMEWHAT LIKELY	NEUTRAL	SOMEWHAT UNLIKELY	VERY UNLIKELY	TOTAL	WEIGHTED AVERAGE
35.87% 33	36.96% 34	20.65% 19	3.26% 3	3.26% 3	92	2.01

The ~10% reduction in the positive response "very likely" between offering a 70% chance of improvement and a 50% chance of improvement is noted. However, FDA does not believe conclusions can or should be drawn from this data. Further, it is not clear if the neuroAD offers a clinically meaningful benefit and if so, how it is able to be quantified as 70% or 50% likelihood.

Appendix VII. Clinical Assessment Scales Commonly Used in AD Therapeutic Trials

Domain / Scale	Description
Cognition	Memory, orientation, language, praxis, etc.
Mini-Mental State Exam (MMSE)	30-pt. scale (higher scores better) Clinician administered patient evaluation Mostly used for eligibility screening and dementia staging
Alzheimer's Disease Assessment Scale-cognitive subscale (ADAS-cog)	70-pt. scale (higher scores worse) Clinician administered patient evaluation Standard cognitive outcome measure in mild-moderate AD
Severe Impairment Battery (SIB)	100-pt. scale (higher scores better) Clinician administered patient evaluation Cognitive outcome measure used in moderate-severe AD
Global Change	Summary outcome assessment from baseline to endpoint
Clinical Global Impression of Change (CGI-C)	7-pt. scale (1 = very much improved, 4 = no change, 7 = very much worse) Clinician rated, based on patient +/- informant interview
Clinician's Interview-Based Impression of Change Plus Caregiver Input (CIBIC-Plus)	7-pt. scale (1 = very much improved, 4 = no change, 7 = very much worse) Clinician rated (with caregiver input), based on semi-structured interview covering cognition, behavior, function
Global Deterioration Scale (GDS)	7-pt. scale (1 = no decline, 7 = very severe decline) Clinician rated based on cognitive change only

Domain / Scale	Description
Function	Activities of daily living (basic and instrumental)
Alzheimer's Disease Cooperative Study Activities of Daily Living (ADCS-ADL)	54-pt. scale (higher scores better) Informant rated interview of 27 basic and instrumental ADL's used in mild – moderate AD; a subgroup of 19 validated items has been used in moderate-severe AD
Disability Assessment for Dementia (DAD)	100-pt. scale (higher scores better) Informant rated interview of 17 basic and 23 instrumental ADL's; initiation, organization, and planning distinguished
Bristol Activities of Daily Living Scale (Bristol ADL)	60-pt. scale (higher scores worse) Informant rated interview of 20 items (10 ADL's, 10 IADL's) each rated on a 0-3 pt. scale
Behavior	Mood , behavior, personality alterations, etc.
Neuropsychiatric Inventory (NPI)	144-pt. scale (higher scores worse) Informant interview of 12 symptom domains rated on a 12-pt. scale based on Frequency (0-4) x Severity (0-3)
Behavioral symptoms in Alzheimer's disease (BEHAVE-AD)	75-pt. scale (higher scores worse) Informant interview of 25 behavioral symptoms rated on a 0-3 pt. scale

Appendix VIII. Statistical Details of Internal Validation of Post-hoc Subgroup

Internal validations generally use cross-validation or bootstrapping in order to discover a predictive subgroup on part of the data and validate its predictive ability on another part of the data. For the internal validation, FDA employed the "Virtual Twin" method (Foster, Taylor, & Ruberg, 2011). At each of 1000 iterations, bootstrapped sampling with replacement from the pivotal data was used to select a training set. Then, a random forest [4] model was fit to each training set, including treatment indicator, baseline ADAS-Cog, as well as products of treatment by baseline ADAS-Cog (Lipkovich, Dmitrienko, & B. R. D' Agostino, 2017). In this ensemble learning method, a split on baseline ADAS-Cog (potentially different at each iteration) was identified as predictive of enhanced success of the Active group over Sham. Then, for each subject in the training set, their predicted response conditional on their being in the Active group as well as in the Sham group was estimated. (One of these predictions will be a "potential" outcome, as each subject was actually only in one of the two treatment groups.) The difference between these two predicted responses is the hypothetical treatment difference for each training subject. The set of treatment differences is used in the second stage of the method as the outcome variables for a regression tree with the goal of identifying a subgroup where the outcome (treatment difference) exceeds -1.20. (The difference of -1.20 was chosen to be small enough so that a split could be identified at all. Otherwise, the regression tree could not identify a split.)

Then, the subjects that were not chosen to be in the bootstrapped sample (the "out of bag" subjects) were used as the test sample for prediction to calculate an adjusted treatment effect (see (Foster et al., 2011) "Method 5 and 6"). The 1000 adjusted treatment effects were averaged. Table 15 shows the result under the label, "Bootstrapped bias-corrected estimate of treatment effect". The "adjustment" is intended to correct for potential spurious values as a result of post-hoc choice. The method is called the "Virtual Twin" method because predicted outcomes are obtained for each training subject on both the Active and Sham (Foster et al., 2011). The average split value on baseline ADAS-Cog was 30.07.

In case there are other predictive covariates in the pivotal study data, we also fit a "Virtual Twin" model that identified a predictive subgroup using other baseline covariates in addition to ADAS-Cog:, e.g., MMSE, gender, age, and whether the subject was taking AD medications at baseline. The simulation process was similar to the method above, but for the fitted random forest and subsequent regression tree, we allowed the model to choose the best predictive splits using all covariates, such that up to 3 splits were chosen (not necessarily on 3 different variables). In general, splits were most often chosen for baseline ADAS-Cog (usually around a value of 30), MMSE (around a value of 22.5), and gender. The adjusted treatment effects were averaged over the 1000 iterations and presented in Table 15 as the "Bootstrapped bias-corrected estimate of treatment effect (using cut-offs on all baseline covariates)".

Standard errors were approximated by repeating the above method 5000 times, each with a bootstrapped sample to represent the entire study data set. The standard deviation of the 5000 adjusted estimates was used to obtain approximate 95% confidence intervals.

---

[4] A random forest regression method averages a large collection of regression trees (e.g., 500) that are (anticipated to be) not highly correlated. Each tree is fit to a bootstrapped sample of the training data, where m covariates are randomly selected from the total number p of covariates (m <= p). For the Virtual Twin method, covariates included a treatment indicator, as well as products of the indicator with covariates (see references in the text for explanations). In the bootstrap method m = 2 covariates were tried at each split of a given tree. A prediction for a (new) subject is made by averaging over the trees, using that subject's covariate values to obtain the predicted leaf node value in each tree. Due to intended low correlations between pairs of trees (from selecting potentially different subjects and different covariates per tree), the variance of the averaged tree is reduced.

References

Bentwich, J., Dobronevsky, E., Aichenbaum, S., Shorer, R., Peretz, R., Khaigrekht, M., . . . Rabey, J. M. (2011). Beneficial effect of repetitive transcranial magnetic stimulation combined with cognitive training for the treatment of Alzheimer's disease: a proof of concept study. *J Neural Transm (Vienna), 118*(3), 463-471. doi:10.1007/s00702-010-0578-1

Chang, C. H., Lane, H. Y., & Lin, C. H. (2018). Brain Stimulation in Alzheimer's Disease. *Front Psychiatry, 9*, 201. doi:10.3389/fpsyt.2018.00201

Foster, J. C., Taylor, J. M., & Ruberg, S. J. (2011). Subgroup identification from randomized clinical trial data. *Stat Med, 30*(24), 2867-2880. doi:10.1002/sim.4322

Heilbronner, R. L., Sweet, J. J., Attix, D. K., Krull, K. R., Henry, G. K., & Hart, R. P. (2010). Official position of the American Academy of Clinical Neuropsychology on serial neuropsychological assessments: the utility and challenges of repeat test administrations in clinical and forensic contexts. *Clin Neuropsychol, 24*(8), 1267-1278. doi:10.1080/13854046.2010.526785

Jacobs, D. M., Ard, M. C., Salmon, D. P., Galasko, D. R., Bondi, M. W., & Edland, S. D. (2017). Potential implications of practice effects in Alzheimer's disease prevention trials. *Alzheimers Dement (N Y), 3*(4), 531-535. doi:10.1016/j.trci.2017.08.010

Lee, J., Choi, B. H., Oh, E., Sohn, E. H., & Lee, A. Y. (2016). Treatment of Alzheimer's Disease with Repetitive Transcranial Magnetic Stimulation Combined with Cognitive Training: A Prospective, Randomized, Double-Blind, Placebo-Controlled Study. *J Clin Neurol, 12*(1), 57-64. doi:10.3988/jcn.2016.12.1.57

Lipkovich, I., Dmitrienko, A., & B. R. D' Agostino, S. (2017). Tutorial in biostatistics: data-driven subgroup identification and analysis in clinical trials. *Stat Med, 36*(1), 136-196. doi:10.1002/sim.7064

Molnar, F. J., Man-Son-Hing, M., & Fergusson, D. (2009). Systematic review of measures of clinical significance employed in randomized controlled trials of drugs for dementia. *J Am Geriatr Soc, 57*(3), 536-546. doi:10.1111/j.1532-5415.2008.02122.x

Nguyen, J. P., Suarez, A., Kemoun, G., Meignier, M., Le Saout, E., Damier, P., . . . Lefaucheur, J. P. (2017). Repetitive transcranial magnetic stimulation combined with cognitive training for the treatment of Alzheimer's disease. *Neurophysiol Clin, 47*(1), 47-53. doi:10.1016/j.neucli.2017.01.001

Pennello, G., & Rothmann, M. (2018). Bayesian Subgroup Analysis with Hierarchical Models. In K. E. Peace, D.-G. Chen, & S. Menon (Eds.), *Biopharmaceutical Applied Statistics Symposium : Volume 2 Biostatistical Analysis of Clinical Trials* (pp. 175-192). Singapore: Springer Singapore.

Piantadosi, S. (1997). *Clinical Trials: A Methodologic Perspective*: Wiley.

Rabey, J. M., & Dobronevsky, E. (2016). Repetitive transcranial magnetic stimulation (rTMS) combined with cognitive training is a safe and effective modality for the treatment of Alzheimer's disease: clinical experience. *J Neural Transm (Vienna), 123*(12), 1449-1455. doi:10.1007/s00702-016-1606-6

Rabey, J. M., Dobronevsky, E., Aichenbaum, S., Gonen, O., Marton, R. G., & Khaigrekht, M. (2013). Repetitive transcranial magnetic stimulation combined with cognitive training is a safe and effective modality for the treatment of Alzheimer's disease: a randomized, double-blind study. *J Neural Transm (Vienna), 120*(5), 813-819. doi:10.1007/s00702-012-0902-z

Schrag, A., Schott, J. M., & Alzheimer's Disease Neuroimaging, I. (2012). What is the clinically relevant change on the ADAS-Cog? *J Neurol Neurosurg Psychiatry, 83*(2), 171-173. doi:10.1136/jnnp-2011-300881

U.S. Food and Drug Administration. (2005). Guidance for the Content of Premarket Submissions for Software Contained in Medical Devices.

U.S. Food and Drug Administration. (2016a). Guidance for Industry and Food and Drug Administration Staff: Applying Human Factors and Usability Engineering to Medical Devices.

U.S. Food and Drug Administration. (2016b). Guidance for Industry and Food and Drug Administration Staff: Factors to Consider When Making Benefit-Risk Determinations in Medical Device Premarket Approval and De Novo Classifications.

Wilcock, G. K., Lilienfeld, S., & Gaens, E. (2000). Efficacy and safety of galantamine in patients with mild to moderate Alzheimer's disease: multicentre randomised controlled trial. Galantamine International-1 Study Group. *BMJ, 321*(7274), 1445-1449.