# Supplementary Materials

# An Evaluation of MSDC-0160, A Prototype mTOT Modulating Insulin Sensitizer, in Patients with Mild Alzheimer's Disease

Raj C. Shah[1], Dawn C. Matthews[2,3], Randolph D. Andrews[2,3], Ana W. Capuano[1], Debra A. Fleischman[1], James T. VanderLugt[4] and Jerry R. Colca[4,*]

[1]*Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, IL, USA;* [2]*Abiant, Inc., Grayslake, IL, USA;* [3]*ADM Diagnostics, LLC, Chicago, IL, USA,* [4]*Metabolic Solutions Development Company, Kalamazoo, MI, USA*

## DETAILED METHODS FOR IMAGING AND COGNITIVE TESTING

### Imaging Acquisition and Analysis

FDG-PET imaging data was obtained at Rush University Medical Center using a Phillips PET/CT Gemini TF 16 slice camera using operating system version and number 4.0.2.145. The site had sFTP capability along with a high-speed network connection for data transmission (IP Address: 144.74.162.203) within 24 hours after scanning. Sinogram data (projection data) was stored in a separate file. Iterative reconstruction for brain imaging was performed. The transmission source was Combined CT and the camera acquired multiple emission acquisitions without changing the bed position (i.e. dynamic acquisitions). Camera quality control (QC) was performed daily on the PET, CT tube warm-up, and CT air calibration. SUV calibration was performed every six months. The site handled radiopharmaceuticals in vials instead of unit doses and had a 10mL or 20mL vial shield (dispensing pig). Intrinsic resolution of the Rush camera was 4.7 mm. The PET scanner had an up-to-date calibration and normalization on the date of each imaging session. A daily QC check was done at the beginning of the day the scanning was completed. The scan was visually inspected for abnormalities. If there was a possibility that the abnormality could impact the PET scan quality, the visit was rescheduled. Daily QC was performed as recommended by Phillips, but typically included a "checkup/calibration" procedure and a water phantom scan. The checkup/calibration procedure guaranteed optimum image quality by warming up the x-ray tube and was performed at startup and within 1 hour prior to any scan. The water phantom was provided quality measurements of 3 parameters. The parameters were the CRT value of water calculated in Hounsfield unites (HU), the pixel noise of images calculated as a standard deviation, and the tube voltages measured directly on the x-ray tubes. These three measurements were determined for all kVp values. Quality control of blood glucose meter was performed according to the manufacturer's or institution's procedure to ensure proper functioning. Quality control of dose calibrator was performed throughout the course of the study. This typically included daily constancy, quarterly linearity, and annual accuracy. No hardware or software upgrades of the PET imaging system occurred during the study duration. The Rush University Medical Center site was appropriately licensed through appropriate state and/or federal agencies to receive and use [18F]-FDG. The site also received both IRB approval and radiation safety committee approval before scanning any subjects.

All participants were screened by study personnel for contraindications to PET scanning. Participants were imaged in the morning and asked to omit all food and fluids (except water) from midnight the night before the scan until after the imaging was completed. Participants scanned later in the day were asked to omit food and fluids (except water) for at least 3 hours prior to the imaging session. When a participant arrived at the imaging center, compliance with the dietary requirement of not having food or drink in less than two hours was confirmed. If a participant ingested food or drink within two hours, participants waited until two hours have elapsed. Once two hours elapsed, blood glucose levels were measured. If blood glucose levels were not less than 180 mg/dL, then waiting an additional amount of time to recheck blood glucose levels was done. Once blood glucose levels were below 180 mg/dL, the participant was asked to use the restroom to empty his/her bladder. Then, the participant was asked to lie comfortably in a bed or reclining chair in a room in which the ambient noise was minimal and the degree of lighting was controlled and minimized. Blankets/pillows were supplied as needed to maximize comfort. Intravenous access using a small butterfly needle or angiocath was done. 185 MBq (5mCI +/- 10%) of [18F]-FDG was drawn and assayed with a dose calibrator and assay time was recorded to the nearest minute. The [18F]-FDG was injected and the syringe and IV line was flushed with 10mL of normal saline. The injection time was recorded to the nearest minute and the IV line was discontinued. The dose syringe then was re-assayed and if the residual activity was 0.1 mCi or greater, the amount was recorded and the amount of the injected dose was corrected for the residual activity. The participant was allowed to rest comfortably in the room with lights dimmed to a level similar to twilight for 20 minutes for incorporation of [18F]-FDG in the brain and the participant's eyes were open (away from any direct light) and the ears remained un-occluded. The participant was directed to maintain their gaze on a constant object, such as the doorknob of the partially closed door to their room, throughout the uptake period. The participant was monitored periodically to be certain of compliance and to ensure that the eyes did not close and the participant remained awake. It was important that no sudden noises or environmental changes occurred within or just outside of the participant's room during [18F]-FDG uptake. All conditions were maintained the same throughout the 2 scanning periods over 12 weeks. At the end of the 20-minute incorporation period, the participant was encouraged to use the restroom to empty their

bladder. Ample time was given to ensure that the participant was on the scanner at 30 minutes after injection and ready for data acquisition to begin. The participant was positioned and secured in the scanner. Time was taken to ensure the participant was properly positioned and could comfortably maintain that position for 30 minutes during the scanning session. Participants removed any bulky items from their pockets and remove eyeglasses, earrings, hair clips/combs, and hearing aids (if possible). The participant was positioned so the head/neck was relaxed which may have involved adding additional pads beneath the neck to provide sufficient support. It was verified that the participant's ears were in a comfortable position, and not pinched. Lasers were used to ensure that there was little or no rotation in either plane. The head was approximately positioned parallel to the imaginary line between the external canthus of the eye and the external auditory meatus. Supportive devices under the back and/or legs were used to help decrease the strain on these regions and prevention of motion in the lower body. Once the participant was positioned, foam pads were placed alongside the head for additional support. Velcro straps and preferably easily removed tape was used to secure the head position. Vacuum bean bags were also used in this process. Participants were offered a "panic button" or reassured that someone was watching or able to hear them at all times. Checking participant positioning and readjusting (if necessary) the position of the participant's head was done often throughout the scanning session. Upon completion, the participant was removed from the scanner and encouraged to void. The participant also was instructed to drink plenty of fluids and to void frequently throughout the day to help reduce radiation exposure.

Procedures similar to ADNI protocols were used in FDG-PET acquisition of images from participants [1]. The Rush Alzheimer's Disease Center at Rush University Medical Center was an active and approved site for the Alzheimer's Disease Neuroimaging Initiative (ADNI, ADNI-GO, and ADNI-2). Rush was approved as a site based on a pair of phantom scans on the 3-D Hoffman brain phantom following a protocol that matched the acquisition and reconstruction parameters. Phantom scans were passed through a quality control process checking for statistical noise, agreement with a digital version of the 3-D Hoffman brain phantom image (the gold standard), and assessed for image resolution and image uniformity. A 30-minute dynamic, 3D scan consisting of six 5-minute frames (preferably fifteen 2-minute frames to allow for additional motion correction) was acquired. All images were corrected using measured attenuation using standard CT acquisition parameters. A check was performed to ensure that the emission and transmission scans were properly aligned before the participant left the imaging session. Images were reconstructed using parameters specific to the scanning system and were reviewed to assess for artifacts and motion. The "raw" PET image sets were converted to a standard Digital Imaging and Communications in Medicine (DICOM) file format using routines. Image file identification followed a standard file identification so that all scans could be easily identified. The file ID was assigned by study personnel prior to the PET scan. All raw and processed study data including copies of the normalization and blank scans were archived. Image data was transferred to Abiant, Inc. (Grayslake, IL), as soon as images had been acquired and reconstructed. All scans were run through a stringent quality control procedure to assess image quality by Abiant, Inc. Quality control checks include evaluation of the following: the number of detected coincidence events (for statistical quality), motion assessment across temporal frames, whether the brain was fully in the field of view, scan artifacts such as asymmetry or streaking, and image header checks to make sure the exact protocol was followed. The different temporal frames were co-registered and placed in a common spatial orientation slice thickness for processing and analysis. Additional spatial normalization and smoothing steps were applied when relevant to the analyses to be performed.

### *FDG-PET Region of Interest Construction.*

The primary analysis of the FDG-PET data was the evaluation of the longitudinal change in regional glucose metabolism (rCMglc) levels from baseline to 12 weeks in the treatment group as compared to the placebo group. Five regions and their composite average were evaluated bilaterally: posterior cingulate, parietal cortex (angular gyrus), lateral temporal cortex, medial temporal cortex, and anterior cingulate-medial frontal cortex. These regions have been reported across a broad set of published work to decline in rCMglc throughout the continuum of Alzheimer's Disease progression and its precursor Mild Cognitive Impairment stages. The pons was used as the primary normalization reference region based on prior work with insulin sensitizers in animal models. Results using cerebellum and whole brain as normalization reference regions were compared to pons-normalized results to verify consistency. ROI sampling was performed by spatially normalizing each participant's FDG scans to a common template and common set of ROIs. The reverse of the transform for each participant's FDG scan was applied to the template ROIs, allowing the participant's original FDG scans to be sampled in their native space without distortion due to spatial normalization or smoothing. The ROIs were developed using an elderly population that includes Mild Cognitive Impairment (MCI) and AD subjects, and their application as described has been validated using larger populations of normal, MCI, and AD subjects from multiple sites and ADNI data.

### *FDG-PET Voxel-Based Assessment*

Voxel-based analyses were performed to identify non-*a priori* clusters that differ within-group between baseline and 12 weeks, and between placebo and treatment groups at baseline and with respect to longitudinal change. Statistical Parametric Mapping was applied to the data set using SPM [2] (Version 5 or later; http://www.fil.ion.ucl.ac.uk/spm/). All scans were spatially normalized to a common anatomical template, and an optimal smoothing filter kernel size was determined by evaluating signal to noise ratios calculated through application of Principal Components Analysis (PCA).

## REFERENCES

[1] Jagust WJ, Bandy D, Chen K, Foster NL, Landau SM, Mathis CA, *et al.* The Alzheimer's Disease Neuroimaging Initiative positron emission tomography core. Alzheimer's & Dementia 6: 221-229 (2010).

[2] Langbuam J.B. Chen K, Lee W, Reschke C, Bandy D, Fleisher AS. *et al.* Categorical and correlational analyses of baseline fluorodeoxy-glucose positron emission tomography images from the Alzheimer's Disease Neuroimaging Initiative (ADNI). Neuroimage 45: 1107-16 (2009).

## Cognitive Testing and Composite Measure Construction

A battery of cognitive function tests was administered in an approximately one hour session. The global cognitive function tests assessed a broad range of dissociable cognitive abilities that appear to have different anatomic substrates commonly affected by aging and Alzheimer's disease and/or widely used for clinical classification of persons with possible dementia. This battery was entirely portable and is administered by trained study personnel. Episodic Memory refers to learning and retention of specific events (episodes) embedded in an autobiographical, temporal context. Semantic Memory refers to previously acquired knowledge independent of the specific context in which the facts were learned. Working Memory refers to the ability to manipulate information held in short-term, limited capacity, memory stores and is an aspect of executive functioning. Perceptual Speed refers to the speed with which perceptual comparisons can be made. Visuospatial ability refers to the ability to process spatial relationships in the visual modality.

The Mini-Mental State Examination [1] was used to describe the cohort but not in analyses, and Complex Ideational Material [2] were used only for diagnostic classification. The remaining 19 tests will represent five domains of cognition. Seven tests assessed episodic memory: Word List Memory, Word List Recall, and Word List Recognition from procedures established by the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) [3]; immediate and delayed recall of Logical Memory Story A from the Logical Memory subtest of the Wechsler Memory Scale—Revised [4]; and immediate and delayed recall of the East Boston Story [5]. Three tests assessed semantic memory: Verbal Fluency,[3] (a 15-item version of the Boston Naming Test [6], subsets of items from Complex Ideational Material [2] and a 15-item reading test [7]. There were three tests of working memory: Digits Forward and Digits Backward from the Wechsler Memory Scale—Revised [4]; and Digit Ordering [8]. Four tests assessed perceptual speed: the oral version of the Symbol Digit Modalities Test, Number Comparison [9], and two indices from a modified version of the Stroop Neuropsychological Screening Test [10]. Finally, there were two tests of visuospatial ability: items from Judgment of Line Orientation [11] and Standard Progressive Matrices [12].

To minimize floor and ceiling artifacts and other sources of measurement error, a summary measure was constructed for global cognition based on 19 tests. The summary measure was constructed by converting the raw scores from the individual tests to z scores, using the mean and standard deviation from the baseline evaluation of all participants in the Rush Memory and Aging Project [13], and averaging the z scores. A summary score was treated as missing if less than half of the component tests had valid scores. Further information on the individual tests and the derivation of the cognitive measures is published elsewhere [14-16].

A 9-item summary measure for executive function was created using the raw scores from Digits Forward and Digits Backward from the Wechsler Memory Scale-Revised, Digit Ordering, two indices from a modified version of the Stroop Neuropsychological Screening Test, Standard Progressive Matrices, Category Fluency Test, Wisconsin Card Sorting test [17], Controlled Oral Word Association [18], and Trail-Making tests [19]. Raw scores from the individual tests were converted to z scores, using the mean and standard deviation from the baseline evaluation of all study participants, and averaging the z-scores. A summary score was treated as missing if less than half of the component tests had valid scores.

The Alzheimer's Disease Assessment Scale-cognitive subscale (ADAS-Cog) is a validated scale for the assessment of cognitive performance and is commonly used in clinical trials of pharmaceutical interventions in Alzheimer's disease [20].

## REFERENCES

[1] Folstein MF, Folstein SE, McHugh PR. "Mini-Mental State": a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 12: 189-198 (1975).
[2] Goodglass H and Kaplan E. The Assessment of Aphasia and Related Disorders (Lea & Febiger, Philadelphia, 1972).
[3] Welsh KA, Butters N, Mohs RC, Beekly D, Edland S, Fillenbaum G, *et al.* The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part V: a normative study of the neuropsychological battery. Neurology 44: 609-614 (1994).
[4] Weschler D. Wechsler Memory Scale – Revised Manual (Psychological Corp, San Antonio, 1987).
[5] Albert M, Smith L, Scherr P, Taylor JO, Evans D, Funkenstei, HH. Use of brief cognitive tests to identify individuals in the community with clinically diagnosed Alzheimer's disease. Int. J. Neurosci. 57: 167-178 (1991).
[6] Kaplan EF, Goodglass H, Weintraub S. The Boston Naming Test. (Lea & Febiger, Philadelphia. 1983).
[7] Nelson HE. National Adult Reading Test (NART) Manual. (NFER-Nelson Publishing Co, Windsor, UK, 1982).
[8] Cooper JA, Sager HJ. Incidental and intentional recall in Parkinson's disease: an account based on diminished attentional resources. J Clin Ex Neuropsychol 15: 713-731 (1993).
[9] Smith, A. Symbol Digit Modalities Test Manual – Revised. (Western Psychological Services, Los Angeles, 1982).
[10] Trenerry, M.R., Crosson, B., DeBoe, J. & Keberm, W.R. The Stroop Neuropsychological Screening Test. Psychological Assessment Resources, Odessa, FL, 1989.
[11] Benton AL Sivan AB, Hamsher K, Varney N, Spreen O. Contributions to Neuropsychological Assessment. 2nd ed. Oxford University Press; New York, 1994.
[12] Raven, J.C., Court, J.H. & Raven, J. Manual for Raven's Progressive Matrices and Vocabulary: Standard Progressive Matrices. Oxford Psychologists Press, Oxford, UK 1992.
[13] Bennett DA, Schneider JA, Buchman AS, Mendes de Leon C., Bienias J, Wilso RS. The Rush Memory and Aging Project: study design and baseline characteristics of the study cohort. Neuroepidemiology 25: 163–175 (2005).
[14] Wilson RS, Barnes LL, Kreuger KR, Hoganson G, Bienias JL. Bennett DA. Early and late life cognitive activity and cognitive systems in old age. J Int Neuropsychol Soc 11, 400-407 (2005).

[15] Bennett DA, Schneider JA, Tang Y, Arnold SE, Wilson RS. The effect of social networks on the relation between Alzheimer's disease pathology and levels of cognitive function in old people: a longitudinal cohort study. Lancet Neurol. 5: 406-412 (2006).
[16] Boyle PA, Wilson RS, Aggarwal NT, Arvanitakis Z, Kelly J, Bienias JL, *et al.* Parkinsonian signs in mild cognitive impairment. Neurology 65: 1901-1906 (2005).
[17] Nelson HA. Modified card sorting test sensitive to frontal lobe defects. Cortex 12: 313-324 (1976).
[18] Benton AL, Hamsher deS. Multilingual Aphasia Examination. AJA Associates, Iowa City, 1989.
[19] Reitan RM. Validity of the Trail Making test as an indicator of organic brain damage. Percept Mot Skills 8: 271-276 (1958).
[20] Rosen WG, Mohs RC Davis KL. A new rating scale for Alzheimer's disease. Am J Psychiatry 141: 1356-1364 (1984).

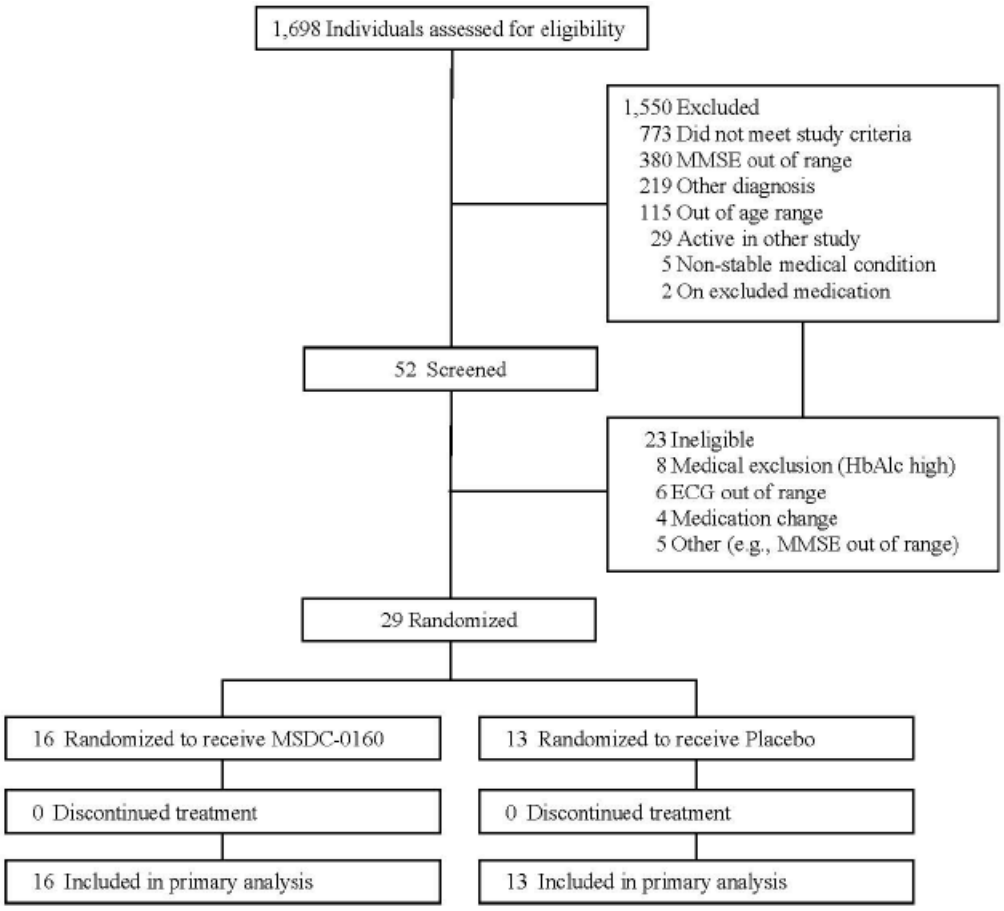


**Fig. (S1).** Study participant flow.

**Table S1. Summary of treatment-emergent adverse events by system organ class and preferred term during the double-blind treatment.**

| System organ class<br>Preferred term | MSDC-0160 150 mg<br>(N = 16)<br>n (%) | Placebo<br>(N = 13)<br>n (%) | Total<br>(N = 29)<br>n (%) |
|---|---|---|---|
| Infections and infestations | 4 (25.0) | 1 (7.7) | 5 (17.2) |
| Urinary tract infection | 3 (18.8) | 0 (0.0) | 3 (10.3) |
| Nasopharyngitis | 0 (0.0) | 1 (7.7) | 1 (3.4) |
| Pneumonia | 1 (6.3) | 0 (0.0) | 1 (3.4) |
| Gastrointestinal disorders | 3 (18.8) | 1 (7.7) | 4 (13.8) |
| Diarrhea | 3 (18.8) | 1 (7.7) | 4 (13.8) |
| Cardiac disorders | 0 (0.0) | 2 (15.4) | 2 (6.9) |
| Patent foramen ovale | 0 (0.0) | 1 (7.7) | 1 (3.4) |

| | | | |
|---|---|---|---|
| QTcB elevation | 0 (0.0) | 1 (7.7) | 1 (3.4) |
| Nervous system disorders | 0 (0.0) | 2 (15.4) | 2 (6.9) |
| Cognitive decline, subjective | 0 (0.0) | 1 (7.7) | 1 (3.4) |
| Syncope, vasovagal | 0 (0.0) | 1 (7.7) | 1 (3.4) |
| General disorders and administration site conditions | 1 (6.3) | 0 (0.0) | 1 (3.4) |
| Edema peripheral | 1 (6.3) | 0 (0.0) | 1 (3.4) |
| Injury, poisoning and procedural complications | 1 (6.3) | 0 (0.0) | 1 (3.4) |
| Skin laceration | 1 (6.3) | 0 (0.0) | 1 (3.4) |
| Metabolism and nutrition disorders | 0 (0.0) | 1 (7.7) | 1 (3.4) |
| Hypoechoic nodule, thyroid lobe | 0 (0.0) | 1 (7.7) | 1 (3.4) |
| Musculoskeletal and connective tissue disorders | 1 (6.3) | 0 (0.0) | 1 (3.4) |
| Pain in hip | 1 (6.3) | 0 (0.0) | 1 (3.4) |
| Psychiatric disorders | 1 (6.3) | 0 (0.0) | 1 (3.4) |
| Depression | 1 (6.3) | 0 (0.0) | 1 (3.4) |
| Reproductive system and breast disorders | 1 (6.3) | 0 (0.0) | 1 (3.4) |
| Scrotal varicose vein rupture | 1 (6.3) | 0 (0.0) | 1 (3.4) |
| Respiratory, thoracic and mediastinal disorders | 0 (0.0) | 1 (7.7) | 1 (3.4) |
| Nose bleed | 0 (0.0) | 1 (7.7) | 1 (3.4) |

**Table S2. Effect of daily MSDC-0160 versus placebo over 12 weeks on weight and bloodliver function test, hemoglobin, and insulin measures.**

| Measure | **MSDC-0160** mean ± SD (n) | **Placebo** mean ± SD (n) | **p-value** * |
|---|---|---|---|
| **Weight, kilograms** | | | |
| Baseline | 71.0 ± 9.5 (16) | 67.7 ± 13.6 (13) | 0.5 |
| Week 12 | 72.2 ± 9.0 (16) | 67.7 ± 13.4 (13) | 0.1 |
| Δ Week 12 - Baseline | 1.3 ± 2.0 (16) | 0.0 ± 1.8 (13) | 0.1 |
| **Alanine aminotransferase (ALT), IU/L** | | | |
| Baseline | 22.2 ± 21.5 (16) | 22.5 ± 7.5 (13) | 1.0 |
| Week 12 | 18.9 ± 7.3 (16) | 21.3 ± 6.9 (13) | 0.4 |
| Δ Week 12 - Baseline | -3.3 ± 20.3 (16) | -1.2 ± 4.6 (13) | 0.7 |
| **Aspartate aminotransferase (AST), U/L** | | | |
| Baseline | 25.4 ± 13.3 (16) | 27.0 ± 4.2 (13) | 0.7 |
| Week 12 | 24.7 ± 5.1 (16) | 26.3 ± 4.2 (13) | 0.4 |
| Δ Week 12 - Baseline | -0.8 ± 13.4 (16) | -0.7 ± 5.4 (13) | 1.0 |
| **Hemglobin, g/dL** | | | |
| Baseline | 13.5 ± 1.3 (16) | 13.8 ± 0.8 (13) | 0.5 |
| Week 12 | 13.0 ± 1.5 (16) | 14.1 ± 0.8 (13) | 0.03 |

| | | | |
|---|---|---|---|
| Δ Week 12 - Baseline | -0.6 ± 0.8 (16) | 0.3 ± 0.7 (13) | 0.004 |
| **Insulin, micro-units/mL** | | | |
| Baseline | 6.5 ± 3.1 (16) | 8.9 ± 5.3 (13) | 0.2 |
| Week 12 | 5.7 ± 2.0 (16) | 8.9 ± 6.4 (13) | 0.1 |
| Δ Week 12 - Baseline | -0.8 ± 3.1 (16) | 0.0 ± 7.6 (13) | 0.8 |

*p-values derived by t-test