**Trial:** Intranasal Insulin for Improving Cognitive Function in Multiple Sclerosis (NCT02988401)
**Author:** Kathryn Fitzgerald
**Statistical Analysis Plan**
**Version 4.0**
**12/9/2022**

I. **Title**: Intranasal Insulin (INI) for Improving Cognitive Function in Multiple Sclerosis
II. **Investigators:** Dr. Ellen Mowry (PI) and Dr. Scott Newsome (PI)
III. **Intervention**
    a. Low dose Intranasal insulin 10 international units BID (LDINI)
    b. High dose Intranasal insulin 20 international units BID (HDINI)
    c. Placebo (PLAC)
IV. **Outcomes:**
    a. Primary outcome
        i. Change from baseline in cognitive function as assessed by the Symbol Digit Modalities Test (SDMT) using raw scores
    b. Secondary outcomes
        i. Adverse events
            1. Adverse events are defined as occurrence or worsening of an undesirable or unintended sign, symptom, or abnormal laboratory test potentially associated with INI. Adverse events for this study will consider: outcomes for which ≥5% experienced. Adverse events will be characterized overall and stratified by treatment arm).
            2. Evaluation of glucose levels (from finger-stick) for first 15 participants following first dose of medication.
            3. Number of adverse events comparing any INI versus placebo or HDINI versus LDINI versus placebo (e.g., two level any exposure to INI versus three level incorporating dosages).
        ii. Change from baseline cognitive function in the tests listed below. Cognitive outcomes will be assessed as raw scores
            1. Change from baseline cognitive function in Controlled Oral Word Association Test (COWAT) assessing fluency (adjusted for form type)
                a. Number of words produced in one minute for each of the three letters (raw score)
                b. Number of words produced in one minute for each of the three letters (adjusted score)
            2. Change from baseline using Benton Judgement of Line Orientation (JLO) test assessing participant ability to match the angle and orientation in space
                a. Total correct
            3. Change from baseline using PASAT-3 test assessing processing speed
                a. Total correct
            4. Change from baseline using Brief Visuospatial Memory Test – Revised (BVLT-R) assessing learning and memory. The following aspects from the BVLT-R will be considered.
                a. Total recall
                b. Delayed recall
            5. Change from baseline using Delis-Kaplan Executive Function System sorting (DKEFS) test for executive functioning. The following aspects from the DKEFS will be considered:
                a. Total description score
                b. Number of correct sorts
            6. Change from baseline using California Verbal Learning Test (CVLT) for learning and memory. The following aspects from the CVLT will be considered and combined into a composite outcome for CVLT using Z scores.
                a. Long-delay free recall correct
        iii. Change in multiple sclerosis (MS) functional composite (MSFC)
            1. Change in components of MSFC. These include walking speed, 9-hole peg test, and PASAT-3.
            2. MSFC will be considered as a Z score; MSFC component tests will be assessed as raw and Z scores.

3. Sustained 20% worsening from baseline MSFC that is confirmed at a subsequent visit.
    a. This test will only be performed if at least 10 events have occurred.

iv. Change in health-related quality of life using the functional assessment in MS (FAMS) scores
    1. These include the total FAMS score and components for symptoms, mobility, fatigue, emotional well-being, social well-being
    2. 1 SD improvement in FAMS total and component scores

v. Change in sleep quality
    1. Total Pittsburgh Quality Sleep Index global score

vi. Change in depression using Beck Depression Inventory-II
    1. Change in total depression score

## V. Descriptive analyses

a. Calculate frequencies of the version of the protocol participants were enrolled.
b. Characterize baseline cohort overall and by randomization arm for demographic, clinical, and patient characteristics in a table with 5 columns (overall, placebo, either INI, HDINI, LDINI) using descriptive statistics based on the variable in question (e.g., means and standard deviations, median and interquartile range, and percentages).
    i. These characteristics will be included: age, sex, race, ethnicity, years of education, ADI (if possible), disease subtype (relapsing remitting MS vs. progressive MS), body mass index (BMI), disease duration, MS disease modifying therapy (DMT; categorized as injectable, oral, infusion/monoclonal ab), number of relapses in the past year, EDSS score, anti-depressant, stimulant use, total number of symptomatic therapies, DXA scan results (total body mass, total fat mass, total lean mass, visceral fat mass, body fat%), version of protocol enrolled under, fasting insulin, fasting glucose, HOMA-IR.
c. Characterize SDMT scores overall and by randomization arm at baseline and over time. Two versions of this analysis will be created: 1) including baseline, and weeks, 6, 12 and 24, 2) including baseline and weeks 6, 12, 24, 36 and 48.
    i. Characterize missing data in SDMT by randomization arm and by time
    ii. Calculate descriptive statistics for SDMT levels by randomization arms and by time.
    iii. Plot means (SD) of SDMT levels by randomization arm over time
    iv. Create spaghetti plots with overlaid loess curves for SDMT by randomization arm over time.
    v. Plot SDMT by month of study enrollment and by randomization arm and time to confirm no differences in allocation by calendar time.

## VI. Outline of protocol changes

Outline of timing of protocol changes by calendar time.

## VII. Consort diagram

a. A consort diagram will be created with the following levels: assessed for eligibility, randomized, treatment allocation, follow-up and analytic cohort.

## VIII. Regression models: Modified intention to treat analyses

a. Analyses will adjust for age, sex, and years of education. We will also incorporate variables with differences between the treatment arms that exceed 10% as identified using results of the descriptive analyses.
b. For analytic models for each outcome, the primary hypothesis test will compare INI (regardless of dose) versus PLAC and will be conducted first. Secondary pairwise tests comparing LDINI versus PLAC, HDINI versus PLAC, HDINI versus LDINI will then be conducted.
c. Analyses will first be conducted first through the week 24 measure (0, 6, 12, 24). Secondary analyses will include weeks 36 and 48 (0, 6, 12, 24, 36, 48) to evaluate a potential sustained effect post cessation of treatment.
    i. The week 24 visit must have occurred within 25 weeks (i.e., the protocol outlines a 1 week visit window).
d. Eligible participants for the primary analysis are those with baseline or screening SDMT data
e. Missing values for key baseline covariates of interest is expected to be small; however, primary analyses will include a missing indicator level for the categorical variable if it is included in analytic models.

f. For mixed effects models, time will be considered as a linear variable in primary analyses.
    i. Sensitivity analyses will evaluate potential non-linearity in changes over time using appropriate methods (e.g., quadratic terms for time or splines).
g. Primary outcome: change in SDMT using raw scores over time
    i. Change in the raw SDMT score over time will be assessed using a linear mixed effects model with unstructured covariance structure, random intercepts. If model fit is improved with the addition of random slopes, they will be added. The term of interest will be the coefficient for the cross-product term between INI and time.
h. Secondary outcomes
    i. Adverse events
    ii. Change from baseline cognitive function in the tests listed below.
        1. Rate of change in each cognitive test will be assessed using a linear mixed effects model with unstructured covariance structure, random intercepts. If model fit is improved with the addition of random slopes, they will be added. The term of interest will be the coefficient for the cross-product term between INI and time.
    iii. Change in multiple sclerosis (MS) functional composite (MSFC)
        1. Rate of change in MSFC Z scores (and component Z scores) using a linear mixed effects model with unstructured covariance structure, random intercepts. If model fit is improved with the addition of random slopes, they will be added. The term of interest will be the coefficient for the cross-product term between INI and time.
        2. Rate of change in raw components using a linear mixed effects model with unstructured covariance structure, random intercepts. If model fit is improved with the addition of random slopes, they will be added. The term of interest will be the coefficient for the cross-product term between INI and time. A log transformation will be applied in if data are skewed based on visual inspection.
        3. For 20% change models, we will fit statistical models if ≥10 events occurred. Time to 20% worsening in MSFC (and components) using a Cox proportional hazards model.
        4. Plot unadjusted results using Kaplan Meier curves for time to event outcomes and test for significance using log-rank test (or other appropriate test if assumptions are not met for log-rank).
    iv. Change in health-related quality of life using the functional assessment in MS (FAMS) score.
        1. Rate of change in quality of life using a linear mixed effects model with unstructured covariance structure, random intercepts. If model fit is improved with the addition of random slopes, they will be added. The term of interest will be the coefficient for the cross-product term between INI and time.
        2. For ≥1 SD improvement, we will fit statistical models if ≥10 events occurred. Time to ≥1 SD improvement in FAMS total score and subscale scores will be assessed via Cox proportional hazards model.
        3. Plot unadjusted results using Kaplan Meier curves and test for significance using log-rank test (or other appropriate test if assumptions are not met for log-rank).
    v. Change in sleep quality
        1. Rate of change in global sleep quality using PSQI using a linear mixed effects model with unstructured covariance structure, random intercepts. If model fit is improved with the addition of random slopes, they will be added. The term of interest will be the coefficient for the cross-product term between INI and time.
    vi. Change in depression using Beck Depression Inventory-II
        1. Rate of change in total depression using a linear mixed effects model with unstructured covariance structure, random intercepts. If model fit is improved with the addition of random slopes, they will be added. The term of interest will be the coefficient for the cross-product term between INI and time.

## IX. Regression models: Planned per protocol analyses

a. For planned per protocol analyses, we will consider the following analyses for the primary outcome raw change in SDMT (or corresponding measures of processing speed)
b. Overview of planned per-protocol analyses

i. Adjusted analyses for additional covariates including DXA-derived body fat, bone density, lean mass, BMI, HOMA-IR
ii. Stratified analyses by pre-shutdown and shutdown/post-shutdown.
    1. Pre-shutdown will be considered as all study visits occurring before March 13, 2020, shutdown will be considered from after March 13, 2020.
iii. Analyses including excluding any screening SDMT for those missing the baseline SDMT.
iv. Stratified analyses by key covariates of interest; these will be tested using an interaction term.
    1. Disability status (EDSS< median, EDSS≥median), history of depression, number of non-MS medications (<5, ≥5), MS subtype (RRMS, progressive), DMT (yes, no)
v. Assessing predictors of drop-out/early censoring
    1. Baseline covariates will be used to assess whether certain characteristics were associated with follow-up time.