Official Title: A Randomized Controlled Trial of Vitamin D Supplementation in Multiple Sclerosis

NCT01490502

March 14, 2022

Statistical Analysis Plan

I. **Title**: The vitamin D to ameliorate multiple sclerosis (VIDAMS) trial: Statistical Analysis Plan

II. **Trial Manuscript Leadership Team:** Ellen M. Mowry, MD, MCR (PI) and Sandra D. Cassard, ScD (Co-I, manuscript lead)

III. **Statistician:** Kathryn C. Fitzgerald, ScD

IV. **Intervention**

a. Low dose vitamin D_3 supplementation (LDVD) of 600 IU/day + glatiramer acetate
b. High dose vitamin D_3 supplementation (HDVD) of 5000 IU/day + glatiramer acetate

V. **Outcomes:**

a. Primary outcome (clinical outcome)
    i. Proportion of participants who experienced a confirmed relapse
b. Secondary outcomes (separated out to clinical and radiographic)
    i. Clinical outcomes:
        1. Confirmed or probable relapse
        2. Annualized relapse rate
        3. Number of relapses requiring treatment
        4. Sustained disability progression
            a. Defined as change in EDSS scores at 1 year that was sustained at year 2. Progression is defined as an increase in EDSS score of 1.0.
                i. Use screening rather than baseline EDSS to compute change if relapse between screening and baseline.
        5. Change in multiple sclerosis (MS) functional composite (MSFC)
            a. Change in components of MSFC. These include walking speed, 9-hole peg test, and PASAT-3.
            b. MSFC will be considered as a Z score; MSFC component tests will be assessed as raw and Z scores.
            c. Sustained 20% worsening from baseline MSFC that is confirmed at a subsequent visit.
        6. Change in low contrast letter acuity (LCLA)
            a. Raw change in number of correct letters
            b. Sustained 20% worsening from baseline LCLA that is confirmed at a subsequent visit.
            c. Sustained change of ≥7 letter (clinically meaningful change)
        7. Change in health-related quality of life using the functional assessment in MS (FAMS) scores
            a. These include the total FAMS score and components for symptoms, mobility, fatigue, emotional well-being, social well-being
            b. 0.5 SD improvement in FAMS total and component scores
        8. Development of hypercalcemia and other adverse events
            a. Hypercalcemia is defined as serum calcium level 0.5 mg/dL above the upper limit of normal; severe hypercalcemia is defined as ≥1 mg/dL above the upper limit of normal.
            b. Self-reported development of kidney stones at any time during follow-up.
    ii. MRI outcomes:
        1. T2 lesion volume
        2. Number of new/enlarged T2 lesions
        3. Number of Gadolinium enhancing (Gd+) lesions
        4. Composite sum of number of new lesions (sum of Gd+ and new/enlarged T2 lesions)
        5. Proportion developing a new lesion over follow-up.
        6. Normalized brain parenchymal fraction (nBPV)
        7. Normalized gray matter volume (nGMV)
        8. Normalized white matter volume (nWMV)

VI. **Descriptive Analyses**

a. Characterize baseline cohort overall and by randomization arm for demographic, clinical, and patient characteristics in a table with 3 columns (overall, LDVD, HDVD) 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, disease subtype (relapsing remitting MS vs. clinically isolated syndrome), body mass index (BMI), previous treatment with a MS disease modifying therapy (DMT), EDSS scores, T2 lesion volume, skin tone, and sun exposure, nBPV, nGMV, and nWMV

b. Characterize 25-hydroxyvitamin D (25[OH]D) levels overall and by randomization arm at baseline and over time (screening, baseline, and months 3, 6, 12, 18, and 24)
   i. Characterize missing data in 25(OH)D levels by randomization arm and by time
   ii. Calculate descriptive statistics for 25(OH)D levels by randomization arms and by time.
   iii. Plot means (SD) of 25(OH)D levels by randomization arm over time
   iv. Create spaghetti plots with overlaid loess curves for 25(OH)D levels by randomization arm over time.
   v. Plot 25(OH)D levels by season (or month) of blood draw and by randomization arm and time to confirm no differences.
   vi. Plot 25(OH)D levels by region of blood draw and by randomization arm and time to confirm no differences.

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: Intention to treat analyses

a. We will incorporate variables with differences between the treatment arms that exceed 10% as identified using results of the descriptive analyses (e.g., sex, race, or ethnicity).
b. Eligible participants for the primary analysis are those with baseline data and at least one follow-up visit.
c. 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. The sensitivity of results will be assessed using pattern mixture models[1] and shared random effects models.[2]
d. For mixed effects models, time will be considered as a linear variable in primary analyses. Secondary analyses will account for time as a categorical variable.
e. Primary outcome: proportion of individuals with a relapse.
   i. Proportion with a confirmed relapse will be assessed using a Cox proportional hazards model to account for follow-up time and accounting for identified potentially unbalanced covariates
   ii. 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).
f. Secondary outcome:
   i. Clinical outcomes
      1. Proportion with a confirmed or probable relapse
         a. Proportion with a confirmed or probable relapse will be assessed using a Cox proportional hazards model to account for follow-up time
      2. Annualized relapse rate (two sets: confirmed, confirmed or probable)
         a. Using an Andersen-Gill model for time to recurrent events correcting for dependence between multiple events assess risk of MS relapse
         b. 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).
      3. Number of relapses requiring treatment
         a. Using a negative binomial model incorporating follow-up time using as an offset term. Results will be presented as rate ratio estimates comparing HDVD vs. LDVD.
      4. Sustained disability progression
         a. Using a Cox proportional hazards model
         b. 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).
      5. Change in multiple sclerosis (MS) functional composite (MSFC)
         a. 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.

b. 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. A log transformation will be applied if data are skewed based on visual inspection.
c. 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.
d. 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).

6. Change in LCLA
   a. Rate of change in LCLA 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.
   b. For 20% change and 7-letter change models, we will fit statistical models if ≥10 events occurred. Time to 20% worsening or 7-letter loss in LCLA will be assessed using a Cox proportional hazards model.
   c. 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).
7. Change in health-related quality of life using the functional assessment in MS (FAMS) scores
   a. 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.
   b. For ≥ 0.5 SD improvement, we will fit statistical models if ≥10 events occurred. Time to ≥ 0.5 SD improvement in FAMS total score and subscale scores will be assessed via Cox proportional hazards model.
   c. 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).
8. Development of hypercalcemia adverse events
   a. Proportion of patients developing hypercalcemia or kidney stones will be compared descriptively
   b. A logistic regression model will be considered if the number of events for either adverse event exceeds 10.

ii. MRI outcomes:

1. T2 lesion volume
   a. Rate of change in T2 lesion volume 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.
2. Number of new/enlarged T2 lesions
   a. Using a negative binomial model incorporating follow-up time using as an offset term.
3. Number of Gadolinium enhancing (Gd+) lesions
   a. Using a negative binomial model incorporating follow-up time using as an offset term.
4. Composite number of new lesions (sum of Gd+ and new/enlarged T2 lesions)
   a. Using a negative binomial model incorporating follow-up time using as an offset term.
5. Proportion developing a new lesion
   a. Proportion developing a new lesion will be assessed using a Cox proportional hazards model to account for follow-up time and accounting for identified potentially unbalanced covariates
   b. 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).
6. nBPV
   a. Rate of change in nBPV 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.

7. nGMV
    a. Rate of change in nGMV 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.
8. nWMV
    a. Rate of change in nWMV 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.

IX. Regression models: Planned per protocol analyses

a. For planned per protocol analyses, we will consider the following analyses for the primary outcome proportion with a confirmed relapse and secondary outcomes including proportion with a confirmed or probable relapse, annualized relapse rate, proportion developing a new lesion, change in overall MSFC Z score.
b. Overview of planned per-protocol analyses
    i. Accounting for change in disease modifying therapy using the following models:
        1. DMTs will be included as a time-varying covariate in Cox models.
        2. Individuals will be censored upon switching DMTs
        3. Inverse probability weighted models account for therapy changes; weights analysis will also incorporate premature censoring (see iv).
    ii. Assessing differences by adherence to study medication or drug
        1. Patient adherence to treatment will be included as a time-varying covariate; the variable will be derived using all available information (missed doses, pill counts).
            a. Likely to categorize to deal with heterogeneity in collection of data
            b. Missing carried forward
        2. Patient adherence to MS DMTs will be included as a time-varying covariate using self-reported patient medication adherence.
    iii. Mediation by change in 25(OH)D levels
        1. The effect of change in 25(OH)D on primary outcomes will be assessed using instrumental variables and 2-stage residual inclusion with study-group assignment as the instrument and change in 25(OH)D level from baseline as the exposure of interest.[3]
    iv. Assessing predictors of drop-out/early censoring
        1. Baseline covariates will be used to assess whether certain characteristics were associated with follow-up time.
        2. Sensitivity analyses will fit inverse probability weighted models accounting for premature censoring or therapy changes (as in i).
    v. Determination of treatment response scores to quantify predictors of participants who will respond "responders"
        1. Treatment response scores will be derived using baseline participant characteristics to assess and quantify potential heterogeneous treatment effects[4–6]
    vi. Assessing whether inclusion of individuals with high inflammatory activity impact the results.
        1. A person with high inflammatory activity is defined as individuals with ≥4 relapses over the study period.
        2. Check the distribution for Gd+ enhancing lesions at baseline and over follow-up. If clear outliers (>+4SD) then analyses excluding these individuals.
    vii. Assessing effect modification by skin tone using Fitzpatrick scale and sun exposure
        1. Stratified analyses will assess effect modification by skin tone and sun exposure.

References

1. Little RJA. Pattern-Mixture Models for Multivariate Incomplete Data. Journal of the American Statistical Association 1993;88(421):125–134.

2. Follmann D, Wu M. An Approximate Generalized Linear Model with Random Effects for Informative Missing Data. Biometrics 1995;51(1):151–168.

3. Terza JV, Basu A, Rathouz PJ. Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling. Journal of Health Economics 2008;27(3):531–543.

4. Bovis F, Carmisciano L, Signori A, et al. Defining responders to therapies by a statistical modeling approach applied to randomized clinical trial data. BMC Med 2019;17(1):113.

5. Zhao L, Tian L, Cai T, et al. EFFECTIVELY SELECTING A TARGET POPULATION FOR A FUTURE COMPARATIVE STUDY. J Am Stat Assoc 2013;108(502):527–539.

6. Pellegrini F, Copetti M, Bovis F, et al. A proof-of-concept application of a novel scoring approach for personalized medicine in multiple sclerosis. Mult Scler 2020;26(9):1064–1073.