

eRADAR: Statistical Analysis Plan (SAP)

Document date: February 27, 2026

NCT #s: NCT05356702 & NCT05905796

Project Title:

Low-Cost Detection of Dementia using Electronic Health Records Data: Validation and Testing of the eRADAR Algorithm in a Pragmatic, Patient-Centered Trial

Purpose of this document:

This document maps out the statistical analysis plan for the eRADAR Aim 2 analyses

Statistical Analyses

Sample update:

For both sites, we will exclude patients who died before baseline i.e. index date (date of initiation) but whose death was not known to the study team at the time of initiation. Dates of death are collected as part of the EHR data pull for outcome analyses, and we may learn then that a patient initiated into the study had actually died for their index date.

At UCSF: Our initial patient identification procedure (at Women's Health) did not exclude patients who have a preferred language other than English, Spanish, Cantonese, and Mandarin; 18 patients with other preferred language were enrolled. **These 18 patients will remain in all analyses.** Patient identification procedures at subsequent clinics excluded patients who did not speak a supported language.

Descriptive analysis:

Descriptive statistics will be summarized for all variables. We will report the mean, standard deviation, median, and IQR for continuous variables and number and percentages for categorical variables. We will summarize all eRADAR variables for preliminary analyses to review internally and report a subset in the primary manuscript; all other variables will be summarized in supplemental tables.

Identifying imbalanced covariates:

Since we have providers as the unit of randomization, patient-level characteristics are less likely to be balanced by randomization (than if randomization had been done on the patient-level). We will compare distribution of baseline covariates between study arms separately for each site using standardized mean differences (SMD). Because regression analyses will adjust for the eRADAR score anyway, we do not need to formally evaluate SMD for the eRADAR score between arms. For predictors included in the eRADAR score, we will first identify any characteristics with a $SMD > 0.1$ and re-evaluate these characteristics within quartiles of the eRADAR score to see if differences persist. (Since regression models will be adjusted for the eRADAR score, we only need to adjust for eRADAR predictors that are not balanced after adjustment.) We note that adjusting for predictors included in the eRADAR score makes interpretation of coefficients associated with these predictors and the eRADAR score non-sensical: one can't interpret the coefficient for age, for example, as the change in the outcome associated with one year increase in age holding all other variables fixed because the eRADAR score would also change with an increase in age. We will also examine the imbalance of covariates not included in the eRADAR score (including race, ethnicity, language, preferred language [UCSF only], number of outpatient visits prior to baseline [categorical variable], and MCI and memory problems). We will adjust for covariates with a $SMD > 0.1$. See meeting notes for detailed description of categories for SMD comparison.

At UCSF, the 18 patients who have a preferred language not supported by the study will be excluded from SMD calculations for language preference due to small sample sizes (but will be included in all

other SMD calculations). If the SMD for preferred language at UCSF is >0.1 (after checking within quartiles of the eRADAR score), then we will adjust for Spanish, Cantonese, Mandarin, and other language (with English as the reference group) in the outcome regression model.

Descriptive analyses will also include summaries of all outcomes, stratified by site and treatment arm. For all the outcomes that were collected as time-to-event outcomes (e.g., time-to-first dementia diagnosis), we will create Kaplan Meier plots illustrating time-to-event.

Safety Reports

The primary approach to safety monitoring will be extraction of group-level data from the EHR for each treatment arm for the following outcomes: new diagnoses of depression, suicide attempts or self-harm, hospitalizations, and deaths. We will include rates of these events by study arm. While the safety report created for DSMB review also included weighted summaries of events for intervention arm participants who completed a brain health visit, with weighting by baseline covariates to account for non-random selection, we will not include these analyses in the manuscript reporting study results due to time constraints. We will not conduct inferential analyses comparing rates of these events between study arms.

Inferential analysis:

Note that:

- Our primary analytic approach will be “intention-to-treat” in which all individuals randomized are included in their original randomization groups, regardless of whether they received the intervention).
- Separate analyses will be performed for each site (KPWA and UCSF).
- The planned analyses (except HTE) will not be performed if the total number of events is <10 per arm.

Health outcomes

Primary analysis – Dementia

We will investigate the impact of eRADAR on new diagnoses of dementia within **12** months of the index date (date of initiation) as assessed from the EHR as our primary outcome. Descriptive analyses will summarize the proportion of patients with new diagnoses of dementia within 12 months of the index date by treatment assignment. For this outcome (binary), we will estimate the relative risk (RR) as the average change in the probability (increase or decrease) of having a new dementia diagnosis for patients in the intervention arm compared to patients in the control arm. Primary unadjusted models will only include binary treatment assignment and eRADAR score. Note: We will adjust for the log-transformed eRADAR score to align with the link function used for the regression model. (See adjusted analyses below for additional covariate adjustment plans.) We will use GEE regression with the log link function, with patient as the unit of analysis, clustering within providers, and robust standard errors (with independent covariance matrix assumed). (A GEE model with log link, clustered observations, and robust sandwich variance estimators is known as a “modified Poisson regression.” This does not assume the outcome is a count variable; robust standard error estimation will obtain accurate variance estimates for the binary outcome.) RRs will be calculated, including a 95% confidence interval (CI) around the estimated RR. If the CI around the RR includes 1, it corresponds to no statistically significant difference in the intervention arm in comparison to the control arm for the primary outcome, new diagnoses of dementia within 12 months of the index date.

- **Primary adjusted analysis** will repeat the analysis with the additional inclusion of any of covariates identified in descriptive analyses as imbalanced between arms. (See section above.)
- **Secondary analyses** will repeat the adjusted analyses described above for dementia diagnosis within 6 and 18 months of the index date.

- **Secondary analyses** will compare time-to-event for dementia diagnosis using follow-up and censoring times collected for each patient. These models will estimate the hazard ratio for dementia diagnosis between patients in the intervention and control arm. We will adjust for provider effects (using cluster() option for coxph function in survival package), eRADAR score, and imbalanced covariates identified for adjustment.
- **Subgroup analyses:** HTE analyses will explore whether eRADAR's impact differs by subgroup (age, gender and race/ethnicity).
 - HTE analyses will include a main effect for subgroup and an interaction between subgroup categories and the intervention.
 - HTE analyses will only include patients with well-defined values for the categories of interest. At both sites, race/ethnicity analyses will only compare intervention effectiveness for people who are Hispanic, non-Hispanic White, non-Hispanic Black, and non-Hispanic Asian. Patients with another race or no recorded race or ethnicity will be excluded from the analysis. At UCSF, gender HTE analysis will exclude the one patient who is non-binary.
 - Only the log-transformed eRADAR score will be included as an adjustment variable in the GEE model (due to low sample size and time constraint to assess for imbalanced covariates within subgroups).
 - We will estimate RRs (and 95% CIs) for each subgroup and display these graphically in a forest plot.
 - We also will use ANOVA to estimate p-values for the statistical significance of the interaction.
 - We will consider performing subgroup analyses for language at UCSF if time permits. These analyses will compare the intervention's effect for people who prefer English, Spanish, Cantonese, and Mandarin, and will exclude people with other language preferences
 - We will conduct these analyses even if there are <10 events in each combination of subgroup and study arm, acknowledging that CIs will likely be wide for smaller subgroups.
- **Stratified analyses (time-permitting)** will be considered for subgroup variables. Due to low sample size in some of the subgroups and time constraint, we will not be performing this analysis. Note: this deviates from the protocol paper which states that we will conduct stratified AND subgroup models that test for interactions.

Notes on secondary outcomes (below):

- Analyses of secondary outcomes will only conduct the main analysis for the intervention effect described above; sensitivity and subgroup analyses will not be repeated for secondary outcomes due to time constraint. (We note that this is a deviation from the protocol paper which stated that stratified and subgroup analyses would also be conducted for secondary outcomes.)
- Analyses of secondary outcomes will be adjusted for the same covariates as the primary outcome unless otherwise noted.

Secondary analysis- MCI or Dementia

We will investigate the impact of eRADAR on new diagnoses of MCI or dementia within **12** months of the index date as assessed from the EHR. For descriptive analysis, we will conduct the same approach as previously stated above (Supplement Table 1).

- The outcome will be defined as a composite binary indicator of new diagnosis of MCI or dementia. MCI will be defined by the ICD-10 code G31.84.
- Analyses will be adjusted for the log-transform eRADAR score, imbalanced covariates (described above), and a for a composite baseline indicator for the presence of amnesic/memory loss diagnoses; late effects of cerebrovascular disease, cognitive deficits

diagnosis; MCI; or degenerative disease of the nervous system, unspecified. Baseline diagnoses are measured in the 2 years prior to patient initiation. (We do not adjust for a more specific baseline indicator of MCI only, as the baseline data were not collected with that level of precision.) As such, the resulting RR estimate can be interpreted as the expected change in the rate of dementia or MCI diagnosis comparing someone in the intervention arm to someone in the control arm with the same baseline status.

- This analysis will use the same GEE approach as described for the primary analysis to estimate RR and 95% CIs.
- A note on earlier analysis plans: We had previously planned to limit this analysis to patients who did not have an MCI diagnosis at baseline (and thus, had already experience the MCI or dementia composite). We have discarded that plan for the approach proposed here for two reasons: 1, we do not have a precise flag for MCI at baseline (the variable we have includes cerebrovascular disease and NOS degenerative disease); 2, excluding those with baseline MCI will not capture the total effect of the intervention, as those patients are most likely to receive a dementia diagnosis. Also note that the prior approach of excluding people with baseline MCI required re-evaluating covariate imbalance in the subset of patients without baseline MCI.

Post-hoc Secondary analysis- broader MCI, ML, or Dementia

A post-hoc secondary analysis will repeat the MCI or Dementia analysis described above with an expanded 12-month composite outcome that also includes amnesic/memory loss diagnoses (R41.3, R41.89); late effects of cerebrovascular disease, cognitive deficits diagnosis (I69.311, I69.11, I69.21, I69.31, I69.81, I69.91); and degenerative disease of the nervous system, unspecified (G31.9). All other aspects of the analysis are as described above for the MCI or Dementia composite analysis (including adjustment variables).

Secondary analysis- Depression

As depression diagnoses were a safety outcome included in study monitoring, not an outcome measuring the intervention's effect or impact on health care utilization, we will not conduct inferential analyses comparing new depression diagnoses following study initiation. Descriptive analyses will summarize new depression diagnoses and other safety outcomes in the intervention and control arms. (We note that this is a deviation from the protocol paper which cited depression diagnoses as a secondary study outcome for which we would conduct inferential analyses.)

Healthcare utilization or burden potentially resulting from eRADAR intervention

Secondary analysis – PCP visits in 6 months

We will investigate the impact of eRADAR on healthcare utilization, measured by the number of PCP visits in the 6 months following the index date as assessed in EHR data. Descriptive analyses will summarize the average number of PCP visits by treatment assignment. As this is a count outcome, we will estimate the relative rate (RR) of PCP visits in the intervention arm compared to the control arm using GEE with the Poisson distribution and the log-link function. Additional GEE specifications for clustering and covariance structure will be the same as the primary analysis; this specification will be robust to overdispersion of the count outcome. We will construct at 95% CI around the estimated RR and compare this CI to the null hypothesis of $RR=1$. Note: while we initially planned to include an offset for the number of data follow-up (under the assumption that a person would only have one PCP visit a day and to account for varying follow-up times), we have removed this aspect of the analysis because patients may have more than one PCP visit in a single day.

Secondary analysis – LABS (TSH, B12, SYP, HIV)

We will investigate the impact of eRADAR on laboratory tests performed within 12 months of the index date as assessed from the EHR. We will define 4 binary outcomes, one for each of Thyroid Stimulating Hormone (TSH), Vitamin B12 (B12), Syphilis and HIV. This analysis will be conducted using the same statistical approach described above for the primary analysis. (Note: Some of these labs may be uncommon; inferential analyses will not be conducted for outcomes with fewer than 10 events per arm)

Secondary analysis – NEU (MRI, CT/CAT Scan)

We will investigate the impact of eRADAR on having neuroimaging test(s) performed within 12 months of the index date as assessed from the EHR. For each of the neuroimaging test(s) performed, we will define 2 binary outcomes: (1) indicator of MRI (2) composite indicator of a brain MRI or CT scan. Descriptive analyses will summarize the number and proportion of patients with each test as well as the composite outcome in each study arm. Inferential analysis of the composite outcome will be conducted using the same approach described above for the primary analysis.

Secondary analysis – SLL or NEU or MAC

We will investigate the impact of eRADAR on having referral(s) or visit(s) for neurology (both sites) or site-specific clinics--Speech, Language and Learning (SLL) at KPWA and the Memory & Aging Clinic (MAC) at UCSF--within 12 months of the index date as assessed from the EHR. For each of neurology and SLL or MAC, we will define two binary outcomes (that is, there will be 4 outcomes total): (1) indicator of a referral and (2) indicator of a visit. An additional note for UCSF: the MAC visit outcome variable should also include MAC e-consults. Descriptive analyses will summarize the number and proportion of patients with each type of referral and visit. Inferential analyses of the outcomes will be conducted using the same approach described above for the primary analysis.

Secondary analysis – Prescription for dementia medication

We will investigate the impact of eRADAR on having any prescription for dementia medications within 12 months of the index date as assessed from the EHR. Medications are identified via order at UCSF and via fill at KPWA. This analysis will be conducted using the same approach described above for the primary analysis.

Healthcare utilization measures that could improve with better detection and support of patients

Secondary analysis – PCP visits in 12 months

We will investigate the impact of eRADAR on healthcare utilization, measured by the number of PCP visits in the 12 months following the index date. This analysis will be conducted using the same approach described in the secondary analysis for PCP visits in 6 months (above)

Secondary analysis – Urgent care and emergency department visits

We will investigate the impact of eRADAR on healthcare utilization measured by the number of days with an urgent care (UC) or emergency department (ED) visits in the 12 months following the index date. Descriptive summaries of the analyses will examine counts of each visit type separately. The outcome to analyze will be a binary indicator of any ED or UC visit in the 12 months following the index date. This analysis will be conducted using the same approach described above for the primary analysis.

Secondary analysis – Inpatient hospitalizations

We will investigate the impact of eRADAR on healthcare utilization with a focus on inpatient (IP) admissions in the 12 months following the index date. We will define two IP outcomes: (1) the number of days hospitalized in the 12 months following the index date and (2) the number of IP admissions in 12 months.

- The number of days outcome will be analyzed with a GEE model with Poisson distribution and log link and including the number of days of follow-up as an offset term (Follow-up is censored at death or health plan disenrollment [KPWA only], and the maximum follow-up for this 12-month outcome is 365 days.).
- The number of IP admissions will follow a similar approach with the exception that no offset will be used; given variable admission length, there is no logical offset for the maximum total number of admissions.
- (To clarify, offsets are used for count outcomes when there is some maximum possible outcome that can't be exceeded. For example, for 365 days follow-up, the maximum number

of days a patient could be hospitalized is 365. The offset-adjusted outcome can be interpreted as a rate or proportion, e.g., IP hospitalization for 10 days out of a year. The benefit of using an offset is that we can account for censoring, or different lengths of follow-up. For example, some patients won't have 365 days follow-up (and thus won't have a max of 365 days hospitalized) due to death or disenrollment. By comparison, when considering the number of hospitalizations in a year, there is no meaningful maximum number of since the length of stay for each hospitalization varies. We will note this limitation and think through the possible implications of this limitation.)

Secondary analysis – No shows

We will investigate the impact of eRADAR on healthcare utilization with a focus on the number of “no shows” for clinic visits in the 12 months following the index date. We will analyze this outcome similarly to other count outcomes, with the exception that no offset will be used; given that patients may have multiple visits on the same day, there is no logical offset for the maximum number of missed clinic visits.