Statistical Analysis Plan

C4HH Statistical Analytic Plan
V 2.26.24

Study Summary

Study description: Our goal is to improve control of cardiovascular (CV) disease risk factors defined by the American Heart Association's Life's Essential 8 (sleep, blood glucose, cholesterol, blood pressure, physical activity, weight, diet, and smoking) by engaging patients experiencing health disparities in an innovative technology-based self-management intervention with linkages to health system providers. Using a patient-level randomized pragmatic trial design, the study findings will provide evidence regarding the best population-based strategy for universal delivery to engage all patients experiencing health disparities in self-management to improve CV disease risk factors.

Setting: We will enroll up to 2,100 patients from three Federally Qualified Health Centers: Denver Health and Hospital Authority, Salud Family Health Centers, or STRIDE Community Health Center.

Patient Population

English/Spanish speaking primary care patients aged 18-89 with:

1. A diagnosis of ≥1 of the following CV risk factors: hypertension, diabetes or hyperlipidemia, and
2. The risk factor is at poor or intermediate health levels as defined by LE8 (e.g., BP≥140/90 mm Hg), and
3. The patient exhibits poor adherence as assessed by each of the enrolling health systems based on available data
4. Patients without an address and cell phone on file; in hospice or palliative care; enrolled in another clinical trial will be excluded

Sites:

- Denver Health and Hospital Authority
    - City and County of Denver's safety-net system that serves 208,000 patients annually. Of these patients, 60% are members of ethnic and racial minority groups.
    - 70% live below the 200% federal poverty level.
- STRIDE Community Health
    - FQHC focused on providing affordable and accessible medical, behavioral health, and dental care among low-income, uninsured, and underserved populations residing outside the City and County of Denver.
    - High burden of CV disease (64.2% hypertension, 81%, hyperlipidemia, and 29.7% diabetes). >95% of patients at >100% of Federal Poverty Level.
- Salud Family Health Centers
    - Established as a small migrant health project to meet the health care needs of farmworkers in rural Colorado, Salud is focused on low-income, medically underserved populations inclusive of the migrant and seasonal farmworker populations.
    - One third of patients at Salud have chronic comorbidities. Salud's patient population is 62% Hispanic.

Eligibility will be identified through patient electronic health record (EHR) data.

Study cohort: We will include patients based on the following: 1) diagnosis of one or more of the following CV risk factors (Table 1); and 2) filled a prescription for medication of interest within the past 100 days (Table 1); and 3) the patient exhibits poor adherence to prescribed medication to treat the CV risk factor; and 4) the risk factor is at poor or intermediate health levels as defined by LE8 (Table 2).

Design

Year 1 (UG3 Phase):
Conduct randomized pilot, enrolling 30 patients from each health care system (ongoing)

Years 2-5 (UH3 Phase):
Conduct pragmatic, patient-level randomized intervention, recruiting patients from the same sites.

- Initial estimate is that > 14k patients will be eligible.
- Identify patient meeting inclusion criteria through the EHR, then send opt out packets. For patients who do not opt-out via return of opt-out postcard, they will need to opt-in to the text messaging campaigns. Patients will have another opportunity to-opt out through texting 'STOP' to the intervention text messages.
- Block randomization (block size of 3)
- Once patients opt-in to the text messaging campaign, the study intervention will be initiated immediately following randomization with delivery of text messages focused on self-management support for the LE8 CV risk factors. Patients will be randomized to one of three study arms: 1) generic text messages; 2) optimized texts with AI chatbot; 3) optimized texts with AI chatbot plus proactive pharmacist management
- Study subjects will be followed for at least 12 months following randomization for the primary outcome (LE8 EHR components measured poor/intermediate at baseline) and for secondary outcomes that includes individual LE8 components, patient self-efficacy, change in Framingham CV risk scores, clinical events (e.g., ER visits and hospitalization), healthcare utilization, and costs over 12 months including outpatient and medications. Subgroup analyses will be performed to identify heterogeneity of treatment effect among the following patient groups of interest: Blacks, Hispanics, rural residents, patients with limited English proficiency, and patients with low-income. Sub-groups will be analyzed for the following outcomes: Framingham CV risk, self-efficacy, and overall LE8.

Intervention

The three primary interventions (study arms) are as follows:

1. **Generic Text** – a series of generic text messages will be delivered to patients focused on one of the Life's Essential 8 topics each week after randomization. The entire LE8 curriculum will be delivered over 9 weeks with the 9th week focused on medication adherence.
2. **Optimized texts with AI chatbot** – a series of text message that includes behavioral nudges will be delivered to patients focused on the Life's Essential 8 topics. Patients can ask questions from the curated chatbot library
3. **Optimized texts with AI chatbot plus proactive pharmacist management** – a series of text message that includes behavioral nudges will be delivered to patients focused on

the Life's Essential 8 topics. Patients can ask questions from the curated chatbot library. Pharmacists will also pro-actively follow-up with patients.

Study Outcomes

Primary and Secondary outcomes

Study subjects will be followed for at least 12 months following randomization for the primary outcome (LE8 EHR components measured poor/intermediate at baseline) and for secondary outcomes that includes individual LE8 components, patient self-efficacy, change in Framingham CV risk scores, clinical events (e.g., ER visits and hospitalization), healthcare utilization, and costs. Subjects who have more than one year of follow-up (up to 3 years depending on when they are enrolled during Years 2-3) will continue to be followed for secondary analyses to assess longer-term outcomes.

The primary outcome will be LE8 overall score.

LE8 Measures: We will enroll patients with the following values of LE8 measures observable in the medical record:

- Lipids: >= 190
- Glucose: HbA1c >= 8.0
- BP: systolic >= 140 or diastolic >= 90

Patients will have one or more qualifying measures from above at baseline to become eligible for the study. Patients will be followed for 12 months, with the LE8 measures repeated at 12 months. The primary outcome will be the LE8 qualifying baseline measures that are also observed at 12 months. If no LE8 qualifying baseline measures are observed at 12 months, the patient will be considered lost to follow-up. Using these qualifying measures from above, we will calculate a baseline LE8 measure (single value of 0-100; mean of qualifying measures at baseline) and a follow-up LE8 measure (single value of 0-100; mean of qualifying measures at follow-up). The difference between baseline and follow-up will be the primary outcome (for LE8).

Secondary outcomes will include:

Patient self-efficacy, change in Framingham CV risk scores, clinical events (e.g., ER visits and hospitalization), healthcare utilization, and costs (outpatient encounters and medications). Subjects who have more than one year of follow-up (up to 3 years depending on when they are enrolled during Years 2-3) will continue to be followed for secondary analyses to assess longer-term outcomes.

Ascertainment of outcomes

Outcomes will be assessed in each patient for 12 months following randomization, using a combination of EHR data and survey data. Subjects with more than 12 months of follow-up will continue to be followed for secondary analyses to assess long-term outcomes.

LE8 Measures: We will enroll patients with the following values of LE8 measures observable in the medical record:

- Lipids: >= 190
- Glucose: HbA1c >= 8.0
- BP: systolic >= 140 or diastolic >= 90

Patients may have one or more qualifying measures from above at baseline to become eligible for the study. Patients will be followed for 12 months, with the LE8 measures repeated at 12 months. The primary outcome will be the LE8 qualifying baseline measures that are also observed at 12 months. If no LE8 qualifying baseline measures are observed at 12 months, the patient will be considered lost to follow-up. Annual measurement of these factors is part of usual

care. Therefore, study pharmacists will place referrals/suggestions for these measures as needed at 12 months for enrolled patients. Not all patients will have measures at exactly 12 months. Therefore, we will take the closest measure in time to the 12-month date to represent the final measure. The window for these final measures will be plus or minus 3 months (9-15 months post baseline). The measure closest in time to the baseline plus 12-month date within that window will be considered the final measure. Using the qualifying measures from above, we will calculate a baseline LE8 measure (single value of 0-100; mean of qualifying measures at baseline) and a follow-up LE8 measure (single value of 0-100; mean of qualifying measures at follow-up). The difference between baseline and follow-up will be the primary outcome (for LE8).

Secondary outcomes: We will use the Self-Efficacy for Managing Chronic Disease 6-item Scale to identify self-efficacy. This scale will be administered at baseline as well as at follow-up (12 months). We will also assess individual components of the LS8 lifestyle factors (e.g., BP control), patient self-efficacy, change in the Framingham risk score for CV disease, clinical outcomes (e.g., CV related hospitalizations), healthcare utilization and costs. We will collect outcome measures via a combination of patient report via text message survey, online survey and/or EHR data. Table 3 below summarizes the patient populations for which these outcomes are relevant.

Several sources of cost data will be used. Inpatient utilization will be measured using diagnostic-related groups (DRGs), outpatient utilization using relative value units (RVUs), and cost of pharmacy utilization using the midpoint between the Federal Supply Schedule (FSS) and the National Average Drug Acquisition Cost (NADAC). Inpatient costs will be estimated by applying national payment weights to DRGs, outpatient costs by applying a national conversion factor to RVUs, and pharmacy costs as the median between the FSS and NADAC. Inpatient costs may not be available for all sites, thus we will focus on outpatient and pharmacy costs as described above.

Statistical analysis plan

Primary analyses

This analysis will be completed consistent with the CONSORT guidelines (http://www.consort-statement.org/) based on the intent to treat principle using all patients randomized. Descriptive analyses will be used to describe the cohort and to check for balance across study arms within strata. The primary outcome LE8 will be calculated during the one-year period following randomization using the EHR. Secondary outcomes will be calculated from the EHR or obtained prospectively from patients by text-message based surveys. We anticipate there will be missing survey and EHR clinical data, which is addressed through our proposed analytic approaches detailed below.

Descriptive analysis

We will use means and standard deviations for continuous variables or counts and proportions for categorical variables. We will describe the following groups of patients, those: eligible, sent an introductory letter, opt out, enrolled at baseline, and complete follow-up. We will describe baseline characteristics of these patient populations. We will use standardized mean differences to assess the balance of patient characteristics across comparison variables, including patients who opt out versus those who enroll, and study arms among those enrolled. As a sensitivity analysis, if any imbalance is detected among enrolled patients, we will adjust for those characteristics in the multivariate modeling approaches detailed below.

Primary hypothesis comparison: There will be a cascade of 2 then 1 comparisons for the primary hypotheses, thus we adjust for multiple comparisons with alpha = (0.5/2). The first two comparisons are below:

- Group 1 (generic text) versus Group 2 (Optimized texts with AI chatbot)
- Group 1 versus Group 3 (Optimized texts with AI chatbot plus proactive pharmacist management)
- If either of these statistical tests are significant, we will make the final comparison:
- Group 2 (Optimized texts with AI chatbot) versus 3 (Optimized texts with AI chatbot plus proactive pharmacist management)

Modeling LE8

We will observe baseline and follow-up values for at least one LE8 measure for all patients enrolled and analyzed, and up to eight LE8 measures. LE8 measures measured poor or intermediate at baseline will be included. The LE8 score is the mean of the sub-components, ranging from 0-100. We will model the change from baseline to follow-up of this score using a general linear mixed models (normal distribution with identity link) with random and fixed effects. Fixed effects will include the baseline LE8 score and treatment arm. A sensitivity analysis will be conducted to additionally adjust for patient characteristics at baseline that remain imbalanced post randomization. Random effects will include intercepts for the health system and potentially assigned study pharmacist.

Modeling Secondary Outcomes

The secondary outcomes including individual LE8 components, patient self-efficacy, change in Framingham CV risk scores, clinical events (e.g., ER visits and hospitalization), healthcare utilization, and costs will be analyzed using similar approaches as the primary outcome with the appropriate models, e.g. Cox survival models for time to clinical event or rehospitalization, generalized gamma regression for cost. For example, the Self-Efficacy for Managing Chronic Disease 6-item Scale results in a score between 0-10, which is the mean of the 6 sub-items. We will model the change from baseline to follow-up of this score using a general linear mixed models (normal distribution with identity link) with random and fixed effects. Fixed effects will include treatment arm and patient characteristics at baseline that remain imbalanced post randomization. Random effects will include intercepts for the health system. Subgroup analyses will be performed to identify heterogeneity of treatment effect among the following patient groups of interest: Blacks, Hispanics, rural residents, patients with limited English proficiency, and patients with low-income. Sub-group analyses will not be adjusted for multiple comparisons as they are exploratory. The results of all sub-group analyses will be presented as a distribution of associations in a single table and figure, and then interpreted appropriately in total.

Treatment Comparisons

We will identify differences between treatment arms using a linear scale for LE8 and secondary outcomes. To identify linear differences for parameters from models, we will implement standardization using counterfactual. This method estimates the expected value of the outcome based on the modeling approaches described above assuming all study participants are exposed to arm 1, then again assuming exposed to arm 2, etc. These estimated outcomes are the basis of the treatment comparison. Uncertainty in these estimates will be quantified through bootstrapping. Covariates will be included in modeling approaches above if/when covariate imbalance is noted in the sensitivity analysis, but no

other statistical variable selection will be performed. The group comparisons are described above. Data will be analyzed using SAS (SAS Institute Inc., Cary, NC) and R software.

Missing data

Patients with missing covariate data will be retained in the study and their missing covariate values imputed using multiple chained equation methods. For patients randomized who later opt-out or drop out, their outcome data will be collected up to the point that they opt-out and will be analyzed along with completers in the primary intent to treat analysis if possible. When outcome data cannot be obtained, every effort will be made to document reasons for these missing observations, and analyses will be carried out as recommended by Little, et al.[153] In particular we will base primary analyses on all observed outcome data and will use estimating equation methods weighted by the inverse probability of the outcome being observed. We will carry out the recommended sensitivity analyses based on pattern mixture models, by assuming various values for difference in means between observed and unobserved data and assessing differences in model conclusions.

Futility Analyses/Stopping Rules:

There will be no pre-specified stopping rules for toxicity of treatment, since the intervention for this trial is minimal risk, consistent with current standards of care, and there are no known adverse reactions to the intervention. All adverse events will be reported to the DSMB.

A futility analysis will be conducted to inform the DSMB on potential futility of continuing the trial. After 50% of patients have been enrolled and completed follow-up, we will conduct futility analyses by estimating the conditional power under three different conditions:
(1) Assume unobserved data follows the same distribution of treatment effects per the planned power analysis.
(2) Assume unobserved data follows the same distribution of treatment effects as observed data to date.
(2) Assume unobserved data follows the treatment effects as observed data to date plus 1 standard error.
These analyses will be presented to the DSMB for consideration for recommendations to stop the trial for futility.

Power Analyses

Sample size calculations using the Life's Essential 8 (LE8):

We conducted a simulation-based power analysis for the LE8 outcome to address feasibility. Using LE8 as the outcome and conservative assumptions about effect size, we will have adequate power (85%) if we analyze 1,677 subjects at the conclusion of the study. If we enroll 2,097 patients and assume a ~20% loss to follow-up, we will still have 85% power to detect a small change in the proportion of patients with an improvement in one of the LE8 risk factors. We define improvement if a patient moves from one category to another as defined by each of the individual LE8 risk factors. For example, the following clinical scenario would be defined as an improvement in a LE8 risk factor: patient starts with non-HDL cholesterol of 200 mg/dL (equivalent to 20 points) and improves to a non-HDL cholesterol of 160 mg/dL (equivalent to 40 points) at the 1-year follow-up.

The above is based on the following assumptions: 1) 25% of the patients in the generic text arm will have an improvement in a LE8 risk factor (lipids, glucose or BP); 2) 28% of patients in the AI chatbot text message group will have an improvement in a LE8 risk factor (3% more than the

generic text group; and 3) 31% of patients in the AI chatbot plus proactive pharmacist support will have an improvement in a LE8 risk factor (6% more than the generic text group). We will have adequate patient populations across the 3 proposed health systems (Denver Health, Salud, and Stride) in which to enroll up to 2,000 patients with 1 or more uncontrolled LE8 risk factor.

Further details of this power analysis are summarized below:

We will target for enrollment patients with the following abnormal values of LE8 measures observable in the medical record:

- Lipids: >= 190
- Glucose: HbA1c >= 8.0
- BP: systolic >= 140 or diastolic >= 90

Using the qualifying measures from above, we will calculate a baseline LE8 measure (single value of 0-100; mean of qualifying measures at baseline) and at follow-up LE8 measure (single value of 0-100; mean of qualifying measures at follow-up). The difference between baseline and follow-up will be the primary outcome (for LE8).

Primary hypothesis comparison: There will be a cascade of 2 then 1 comparisons for the primary hypotheses, thus we adjust for multiple comparisons with alpha = (0.5/2). The first two comparisons are below:

- Group 1 (generic text) versus Group 2 (Optimized texts with AI chatbot)
- Group 1 versus Group 3 (Optimized texts with AI chatbot plus proactive pharmacist management)
- If either of these statistical tests are significant, we will make the final comparison:
- Group 2 (Optimized texts with AI chatbot) versus 3 (Optimized texts with AI chatbot plus proactive pharmacist management)

Small (conservative) effect size assumptions:

- All groups: patients have a 5% chance of decreasing one LE8 category on a qualifying health state category.
- Group 1: patients have a 25% chance of increasing one LE8 category on a qualifying health state category.
- Group 2: patients have a 28% chance of increasing one LE8 category on a qualifying health state category.
- Group 3: patients have a 31% chance of increasing one LE8 category on a qualifying health state category.

Moderate effect size assumptions:

- All groups: patients have a 5% chance of decreasing one LE8 category on a qualifying health state category.
- Group 1: patients have a 25% chance of increasing one LE8 category on a qualifying health state category.
- Group 2: patients have a 30% chance of increasing one LE8 category on a qualifying health state category.
- Group 3: patients have a 35% chance of increasing one LE8 category on a qualifying health state category.

The power and sample size estimates are summarized in table 4 below.

Tables

Table 1: CV risk factors and medication classes for eligibility criteria	
Condition	**Classes of medications**
Hypertension	Beta-blockers (B-blockers), Calcium Channel Blocker (CCB), Angiotensin converting enzyme inhibitors (ACEi), Angiotensin Receptor Blockers (ARB), Thiazide diuretic
Hyperlipidemia	Statins, Ezetimibe
Diabetes	Alpha-glucosidase inhibitors, Biguanides, DPP-4 inhibitors, Sodium glucose transport inhibitor, Meglitinides, Sulfonylureas, Thiazolidinediones, and statins

Table 2: Definition and scoring approach for the American Heart Association's Life's Essential 8 score. (PMID: 36775706)		
Blood lipids	Measurement: Plasma total and HDL cholesterol with calculation of non–HDL cholesterol Example tools for measurement: Fasting or non-fasting blood sample	Metric: Non–HDL cholesterol (mg/dL) Scoring: PointsLevel 100<130 60130–159 Eligibility for those with hyperlipidemia (might change based on pilot data) 40160–189 20190–219 0≥220 If drug-treated level, subtract 20 points
Blood glucose	Measurement: FBG or casual HbA1c Example tools for measurement: Fasting (FBG, HbA1c) or non-fasting (HbA1c) blood sample	Metric: FBG (mg/dL) or HbA1c (%) Scoring: PointsLevel 100No history of diabetes and FBG <100 (or HbA1c <5.7) 60No diabetes and FBG 100–125 (or HbA1c 5.7–6.4) (prediabetes) 40Diabetes with HbA1c <7.0 Eligibility for those with diabetes (might change based on pilot data) 30 Diabetes with HbA1c 7.0–7.9 20 Diabetes with HbA1c 8.0–8.9 10 Diabetes with Hb A1c 9.0–9.9

		0 Diabetes with HbA1c ≥10.0
BP	Measurement: Appropriately measured systolic and diastolic BPs Example tools for measurement: Appropriately sized BP cuff	Metric: Systolic and diastolic BPs (mm Hg) Scoring: PointsLevel 100<120/<80 (optimal) 75120–129/<80 (elevated) Eligibility for those with hypertension(might change based on pilot data) 50130–139 or 80–89 (stage 1 hypertension) 25140–159 or 90–99 0≥160 or ≥100 Subtract 20 points if treated level

Table 3: Outcomes and relevant patient populations.	
Outcome	Patient population based on the presence of specific comorbidity for which the outcome is relevant
Outpatient Blood Pressure	Hypertension
LDL/HDL/Total	Hyperlipidemia
FBG/HbA1c	Diabetes
Framingham Risk Score	All
All Cause Hospitalization (1 Yr.)	All
Cause Specific Hospitalization	
Hypertension Emergency	Hypertension
Myocardial infarction (MI)	HTN/Hyperlipidemia/Diab
Stroke	HTN/Hyperlipidemia/Diab

Heart Failure	HTN/
Hyperglycemia	Diabetes
All Cause ED Visit (1 Yr)	All
Cause Specific ED Visit	
Hypertension Emergency	Hypertension
MI	HTN/Hyperlipidemia/Diab
Stroke	HTN/Hyperlipidemia/Diab
Heart Failure	HTN/
Hyperglycemia	Diabetes
Procedures	
PCI	HTN/Hyperlipidemia/Diab
CABG	HTN/Hyperlipidemia/Diab

Table 4: Power and Sample Size estimates

Effect Size	Power	Total sample size assuming 20% dropout	Total sample size assuming 10% dropout	Total Sample size across 3 arms (number analyzed)	Sample in each arm (number analyzed)
Moderate	80%	1,815	1,614	1,452	484
Moderate	85%	2,097	1,864	1,677	559
Small (conservative)	80%	4,665	4,147	3,732	1,244
Small (conservative)	85%	5,157	4,584	4,125	1,375

References

Lumley T, Diehr P, Emerson S, Chen L. (2002) The importance of the normality assumption in large public health data sets. Annual Review of Public Health 23:151-169.

Bell ML, Horton NJ, Dhillon HM, Bray VJ, Vardy J. (2018) Using generalized estimating equations and extensions in randomized trials with missing longitudinal patient reported outcome data. Psycho-Oncology, 27:2125-2131.

Dmitrienko A, Tamhane AC, Wiens BL. (2008) General multistage gatekeeping procedures. Biometrical Journal 50: 667-677.

Fiero MH, Hsu C-H, Bell ML. (2017) A pattern-mixture model approach for handling missing continuous outcome data in longitudinal cluster randomized trials. Statistics in Medicine 36:4094-4105.

White IR, Horton NJ, Carpenter J, Pocock SJ. (2011) Strategy for intention to treat analysis in randomized trials with missing outcome data. British Medical Journal 342:d40.

Hernán, M. A., Brumback, B. A., & Robins, J. M. (2002). Estimating the causal effect of zidovudine on CD4 count with a marginal structural model for repeated measures. Statistics in medicine, 21(12), 1689-1709.

Preisser, J.S., Lohman, K.K. and Rathouz, P.J., 2002. Performance of weighted estimating equations for longitudinal binary data with drop-outs missing at random. Statistics in medicine, 21(20), pp.3035-3054.

Robins, J.M., Rotnitzky, A. and Zhao, L.P., 1995. Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. Journal of the American Statistical Association, 90(429), pp.106-121.

Lin G & Rodriguez RN. (2015) Weighted methods for analyzing missing data with the GEE procedure. SAS paper 166-2015, https://support.sas.com/resources/papers/proceedings14/SAS166-2014.pdf.

Xu C, Li Z, Wang M (2018). wgeesel: Weighted Generalized Estimating Equations and Model Selection. R package version 1.5, https://CRAN.R-project.org/package=wgeesel.