Statistical Analysis Plan

The SPRING Trial: Comparing Suicide Prevention Interventions to Guide Follow-up Care

Comparing the Effectiveness of Two-Way Caring Contacts Texts vs One-Way Caring Contacts Texts vs Enhanced Usual Care to Reduce Suicidal Behavior in Youth and Adults Screening At-Risk for Suicide in Primary Care or Behavioral Health Clinics: The SPRING Trial

Principal Investigator

Anna Radin, DrPH, MPH, Applied Research Scientist, St. Luke's Health System; 208-381-8468; radina@slhs.org

Statistical Analysis Plan prepared by:
Siobhan Brown, PhD (University of Washington)

~~Version 1.0, January 12, 2024~~
Version 2.0, February 27, 2025

1 INTRODUCTION

Suicide is a leading cause of death in the United States; Idaho's suicide rate is consistently among the top ten nationally, and in 2020 was the fifth highest in the US, 72% higher than the national average.[1] The American Foundation for Suicide Prevention (AFSP) has called for research to identify effective interventions that, if implemented at scale, could reduce suicide deaths by 20% by 2025.5 Caring Contacts is one of the few evidence-based interventions shown to reduce suicide deaths,[2,3] and it is feasible to implement at scale, including in rural and low-resource settings.[4-6] The Caring Contacts model involves sending brief, non-demanding expressions of care to suicidal individuals, to let them know that they are being thought of and as a gentle reminder that help is available if needed. The first Caring Contacts efficacy trial randomized patients with recent suicide attempts who refused ongoing care to receive either typed caring letters or no further contact. In other randomized controlled trials, Caring Contacts have proven effective when delivered by post cards,[7,8] and more recently, two-way text messages.[9] A recent trial of text message-based Caring Contacts found that compared to usual care, Caring Contacts were associated with lower odds of experiencing any suicidal ideation since baseline and lower odds of suicide attempt.

2 SPECIFIC AIMS

- **Aim 1:** Measure the effectiveness of augmenting best available usual care with two-way Caring Contacts text messages (CC2) and one-way Caring Contacts texts (CC1) compared to best available usual care alone (UC) for preventing suicidal behavior.
    - **Aim 1 Hypothesis:** We hypothesize under the alternative that CC2 and CC1 are more effective than usual care.

- **Aim 2:** Determine whether CC1 are noninferior to CC2 for preventing suicidal behavior.

- **Aim 2 Hypothesis:** We hypothesize under the null that CC1 are inferior to CC2.

- **Aim 3:** Describe the feasibility of implementing CC2 and CC1 in partnership with a state crisis and suicide prevention hotline.
    - Hypothesis: CC2 will require more resources than CC1 but both will be acceptable and feasible to deliver through a state crisis and suicide hotline.

3 DESIGN

3.1 The proposed study is a pragmatic randomized controlled trial with participants assigned to one of three intervention arms: two-way caring text messages (CC2), one-way caring text messages (CC1), and best available usual care alone (UC). Randomization will occur at the individual level and will be stratified by age (12-17 vs 18+ years). Participants will be randomized in a 1:1:1 ratio to one of the three intervention arms. Randomization will occur prior to enrollment during the informed consent process so that participants can see the intervention arm to which they were randomized before they decide whether to enroll in the study, but remain masked to the other treatment conditions. This trial will be single masked, with most members of the study team including the senior statistician masked to aggregate data by treatment arm. Masking interventionists or participants to the assigned intervention is not feasible due to the nature of the intervention.

3.2 Power & Sample Size

Because the study is designed to assess both the effectiveness of Caring Contacts as compared to usual care as well as the noninferiority of CC1 as compared to CC2, there are two components to the sample size and power calculations. Both assume a standard deviation of 5 in the area under the curve (AUC) of HASS Active Suicidal Behavior Subscale scores over 12

months based on previously published studies[16]and data from an ongoing trial with suicidal patients at SLHS.[4] With a sample size of 759, we will have 80% power to detect a true difference between the CC arms and the UC arm of 1.26 units in the AUC of HASS scores over 12 months accounting for one formal interim analysis for efficacy. Similarly, a sample size of 759 will allow 89% power to determine noninferiority of CC1 to CC2 if the two interventions are truly equivalent, when using a noninferiority margin of 1.0. We plan to enroll 849 participants to account for up to 30% potential lost to follow-up.

Table 1. Power and Sample Size Calculations

Power for superiority

Improvement in CC arm compared to UC arm	Power	Sample Size per Arm	Total Sample Size
1.26	80%	253	759

Power for noninferiority

Noninferiority Margin	True difference between CC1 and CC2	Power	Sample Size per Arm	Total Sample Size
-1.0	0	89%	253	759

3.3 Planned Interim Analyses

An interim analysis is planned for when 50% of 6-month outcome data are available (or as determined by the Data Safety and Monitoring Board) to assess grounds for early stopping if one of the arms is overwhelmingly more effective than the other, and/or to re-calculate sample size if the observed data depart from the assumptions used for the initial sample size calculations.

With input from the DSMB, prior to initiation of the trial, we will finalize a monitoring plan to guide dropping an arm from or early termination of the study. Factors influencing

stopping decision may include (a) formal stopping rules based upon the primary analysis, (b) information on safety outcomes by treatment group, (c) consistency between results for primary and secondary outcomes, and (d) consistency of treatment effects across subgroups. The proposed formal stopping boundaries will be symmetric, two-sided designs which are included in the unified family of group sequential stopping rules.[10,11] Point estimates from interim analyses will be based on the bias adjusted point estimate.[12] Confidence intervals and p-values will be calculated from the ordering of the outcome space based upon the maximum likelihood estimate.[13] The boundaries will be applied to each of the two-way comparisons between the three treatment arms.

Table 2. Interim Analysis Stopping Boundaries

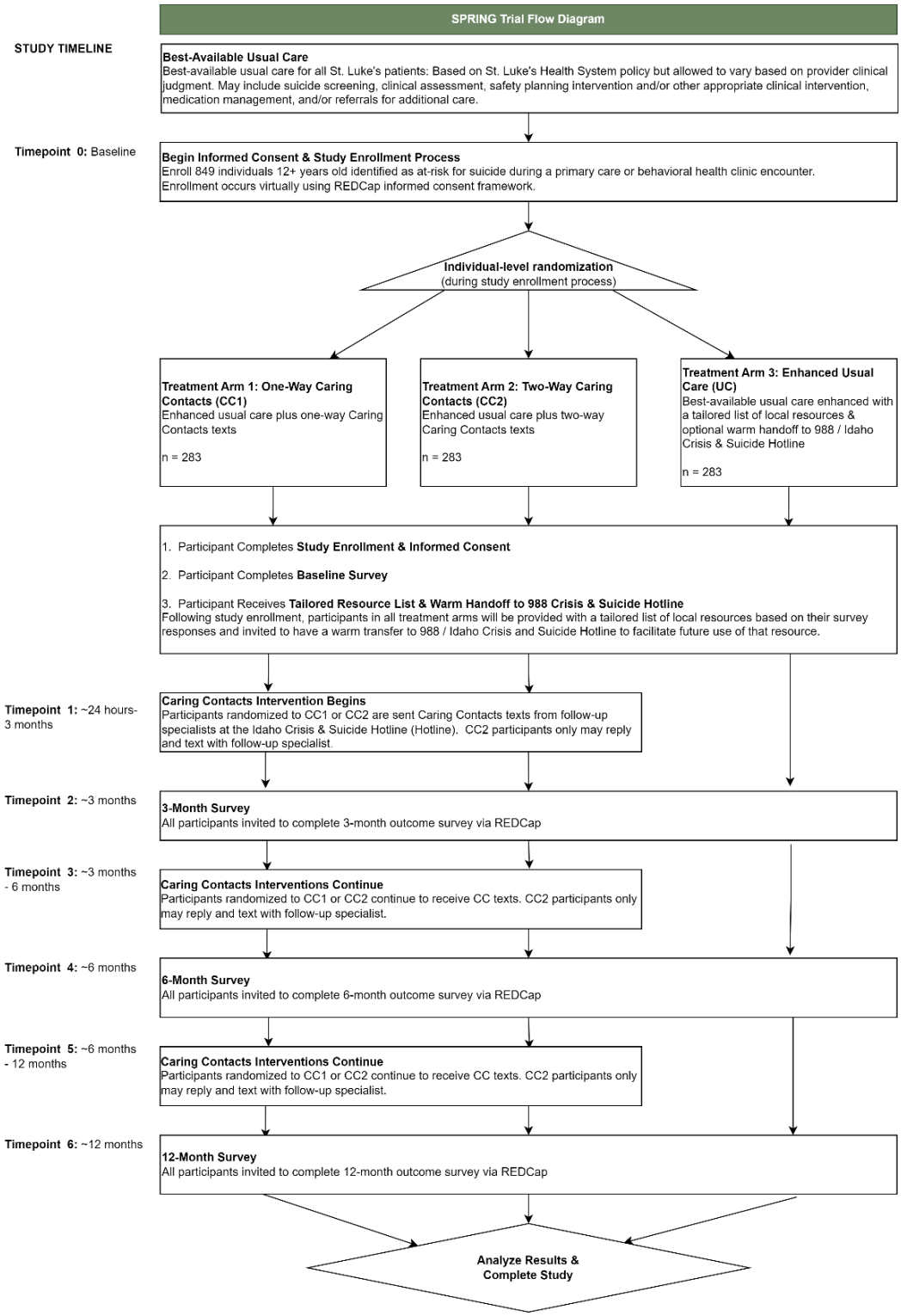
			Lower Stopping Boundary			
Interim Analysis	**Cum. Sample Size with 6 Month Outcome**	**Prop. Max Stat Info**	**Absolute Difference**	**Adjusted Difference**	**95% Confidence Interval**	**P-value**
1	380	0.50	-1.56	-1.44	(-2.54, -0.20)	0.013
2	759	1.00	-0.90	-0.83	(-1.76, 0.00)	0.043

			Upper Stopping Boundary			
Interim Analysis	**Cum. Sample Size with 6 Month Outcome**	**Prop. Max Stat Info**	**Absolute Difference**	**Adjusted Difference**	**95% Confidence Interval**	**P-value**
1	253	0.33	1.56	1.44	(0.20, 2.54)	0.013
2	506	0.67	0.90	0.83	(0.00, 1.76)	0.043

Note: sample sizes given are the cumulative number of study participants with the six-month outcome available at the time of analysis; we expect about 2/3 of that number will be available for the comparison of each two study arms. These calculations assume a standard deviation of 5; at the time of the analysis, the estimated standard deviation from the study data will be used. The boundaries used will be adjusted accordingly. The schedule of the interim analyses may be tweaked slightly based on the DSMB schedule. The information fraction will be updated to account for these differences as well as

participants with partial outcomes available.[14] The p-value boundaries are nominal two-sided p-values.

Figure 1: SPRING Trial Flow Diagram



4 DATA SOURCE

Table 3: SPRING Trial Variables of Interest and Other Data Elements

Variable	Tool/Source	Routinely collected as usual care	Mode of contact	Who will collect
Primary, Secondary, Safety & Exploratory Outcomes (Baseline + 3, 6, 12 months)				
Primary Outcome				
Suicidal behavior	HASS Active Suicidal Behavior Sub-Scale (all surveys)	No	REDCap survey	SLHS
Secondary Outcomes				
Suicidal ideation	HASS Passive Suicidal Ideation Sub-Scale (all surveys)	No	REDCap survey	SLHS
Suicide attempts	C-SSRS (all surveys)	No	REDCap survey	SLHS
Outpatient mental health treatment	Self-report, Epic, Claims data (all surveys)	Yes	REDCap survey	SLHS
Use of crisis care (ED visits, hospitalizations) for suicidality	Self-report, Epic, Claims data (all surveys)	Yes	REDCap survey	SLHS
Safety & Exploratory Outcomes				
Perceived burdensomeness & thwarted belongingness	INQ-15 (all surveys)	No	REDCap survey	SLHS
Perceived mattering to others	GMS (all surveys)	No	REDCap survey	SLHS
Non-suicidal self-injury (NSSI)	SITBI-R (all surveys)	No	REDCap survey	SLHS
Suicidal ideation	SITBI-R (all surveys)	No	REDCap survey	SLHS
Suicidal ideation & behavior	C-SSRS (all surveys)	No	REDCap survey	SLHS
Suicide Plans & Preparatory Acts	SITBI-R (all surveys)	No	REDCap survey	SLHS
Suicide attempts	HASS Suicide Attempts Sub-Scale (all surveys) and SITBI-R (all surveys)	No	REDCap survey	SLHS
Suicide deaths including cause of death	Vital records, Epic	Yes	N/A	Idaho Dept of Health & Welfare, SLHS
All-cause mortality including manner & cause of death	Vital records, Epic	Yes	N/A	Idaho Dept of Health & Welfare, SLHS
ED utilization or hospitalization (all-cause), including number and diagnoses/reason for visit	All surveys, Epic	Yes	REDCap survey	SLHS
Current Suicidal Crisis	All surveys	No	REDCap survey	SLHS
Use of 988 or suicide hotline	All surveys	No	REDCap survey	SLHS
Sociodemographic & Other Exposure Variables				

Age in years at enrollment	Epic, Informed Consent Form	Yes	REDCap survey	SLHS
Sex assigned on birth certificate	Epic	Yes	n/a	SLHS
Race and ethnicity	Epic	Yes	n/a	SLHS
Address including Zip code of residence (urban/rural)	Epic	Yes	n/a	SLHS
County of residence (urban/rural)	Epic	Yes	n/a	SLHS
Gender identity, pronouns, transgender status, & sexual orientation	Baseline survey	No	REDCap survey	SLHS
Marijuana and illicit drug use	Baseline survey, 3, 6, and 12-month outcome surveys	No	REDCap survey	SLHS
Alcohol use	Baseline survey, 3, 6, and 12-month outcome surveys	No	REDCap survey	SLHS
Religion & religious practice	Baseline survey	No	REDCap survey	SLHS
Military / veteran status	Baseline survey	No	REDCap survey	SLHS
Socioeconomic status (employment, education / maternal education, housing stability, income, food/income security)	Baseline survey	No	REDCap survey	SLHS
Depression	Baseline survey (PHQ-A / PHQ-9)	No	REDCap survey	SLHS
History of suicide attempts & self-harm; lethal means	Baseline survey	No	REDCap survey	SLHS
Insurance Provider	Epic	Yes	n/a	SLHS
Suicidal ideation & behavior at referring encounter (usual care)	C-SSRS (in Epic)	Yes	n/a	SLHS
Quality of life at baseline	Baseline survey Euro-Qol	No	REDCap survey	SLHS
Overall health (presence of co-morbidities) at baseline	Epic	Yes	n/a	SLHS
Treatment history and medication use	Epic, REDCap surveys	Yes	REDCap survey	SLHS
Referring clinic name & referring provider	Epic	Yes	n/a	SLHS
Intervention assignment (CC2, CC1, UC)	REDCap	No	n/a	SLHS
Patient Satisfaction				
Patient satisfaction with study intervention	Outcome surveys	No	REDCap survey	SLHS
Process Variables				
Screening rates & results by clinic and provider; referral rates by clinic & provider	Epic / D&A dashboard	Yes	n/a	SLHS

Safety Planning/Connection & Support Planning completion	Epic / D&A dashboard / REDCap / Mosio	No	n/a	SLHS
Individual who completed safety plan	Epic / REDCap / Mosio	No	n/a	SLHS
"Dose" of follow-up contact: timing, type (phone vs text), and number of attempted and successful contacts from the Hotline; # outgoing & incoming texts + content of texts	Mosio	No	n/a	Hotline
Other Variables				
Medical record number (MRN) & encounter number (CSN)	Epic	Yes	n/a	SLHS
Encounter date(s) and time(s)	Epic	Yes	n/a	SLHS
Clinic specialty/type	Epic	Yes	n/a	SLHS
Provider(s)	Epic	Yes	n/a	SLHS
Referral(s)	Epic	Yes	n/a	SLHS
Insurance type	Epic	Yes	n/a	SLHS
Does cell phone on record belong exclusively to study participant or is it shared?	Baseline survey	No	REDCap survey	SLHS
Alternative modes of contact	Baseline survey	No	REDCap survey	SLHS
Death (manner and cause), including suicide deaths and all-cause mortality	Epic and vital records	Yes	n/a	SLHS & Idaho Department of Health & Welfare

4.1 Data Management

Data management will occur using the Research Electronic Data Capture (REDCap) tool, a HIPAA compliant web-based research application used for building and managing online surveys and providing a secure electronic database. Our instance of REDCap is owned and managed by the University of Washington's Institute of Translational Health Sciences. REDCap will be used as the central location for online study data storage and participant management. The database is safeguarded against unauthorized access by established security procedures; appropriate backup copies of the database and related software files will be maintained. The statistics team will compile data from REDCap on a weekly basis for reports, as well as to build and maintain a complete dataset. Baseline data from participants' electronic health records will be entered via REDCap data extraction forms. Twilio, a HIPAA-compliant texting platform integrated into REDCap, will be used to send survey links and survey reminders to participants. Mosio,[15] a HIPAA-compliant online texting and calling platform, will be used to deliver Caring Contacts, to schedule enrollment calls, to send retention texts and survey reminders, and to make phone calls. Mosio records any attempt at contact and successful contact made. To protect the confidentiality of participants, data and associated documentation will be available to approved study personnel only, under a data-sharing agreement that includes a commitment to: (1) use the data only for research purposes; (2) secure the data using appropriate technology; and (3) destroy or return the data after analyses are completed.

5 ANALYSIS SETS/POPULATIONS/SUBGROUPS

5.1 Inclusion & Exclusion Criteria

5.1.1 Study Participant Inclusion & Exclusion Criteria

Inclusion criteria are intentionally broad, to facilitate recruitment of a study sample that is maximally representative of the population of patients that SLHS serves and to maximize external validity of the results.

Inclusion Criteria

- Adolescents (12-17 years old) and adults (18+)
- Response of "yes" to at least one item on the Columbia Suicide Severity Rating Scale (C-SSRS) six-item screener at a SLHS primary care or behavioral health clinic, or electronic health record or provider note from an eligible encounter indicates suicide risk
- Ability to send and receive text messages
- Ability to receive phone calls
- Ability to receive emails
- Participant and legal guardian (if applicable) speak, read, and understand English
- Accommodations may be made for individuals with impaired hearing

Exclusion Criteria

- Individuals who participated in a previous SLHS randomized controlled trial related to Caring Contacts (SPARC Trial or MHAPPS Trial).
- Patients who are unable or unwilling to provide informed consent*, for example, due to acute or chronic cognitive impairment (i.e.: acute psychosis, intoxication, or intellectual disability).
- Primary Care Provider, Behavioral Health Provider, or Principal Investigator determines that participation in the research is not in the best interest of the patient or the study team.

5.2 Analysis Populations

Analyses will be completed using the following populations:

1. **Intention to Treat (ITT) Analytic Population**: Data for all participants that complete study enrollment and the baseline survey will be included in this dataset. Participants will be grouped according to their randomized treatment arm, regardless of the intervention received.
2. **Safety Analysis Dataset**: The safety analysis will include data for all participants who completed study enrollment and the baseline survey (e.g., the ITT Analytic Population Dataset).
3. **Per-Protocol Analysis Dataset**: Data for a subset of participants who were retained for the duration of the study and received the assigned intervention. For participants randomized to either of the caring contacts study arms, this means receiving all 25 scheduled messages. Additionally, participants who receive any of the interventions that are not part of their randomized study arm will be excluded from the per-protocol analyses (e.g. a participant who is randomized to usual care but accidentally receives a caring contact message will be excluded).
4. **Additional Datasets**: Additional datasets may be developed to complete sensitivity analyses, for example, where missing data have been imputed using different techniques.

5.3 Subgroups

Subgroup analyses will be completed based on:

- age at enrollment (12-17; 18-24; 25-49; 50+),
- sex at birth (female vs. male),
- gender identity (cisgender vs. transgender or gender nonconforming),
- sexual orientation (heterosexual vs. homosexual, bisexual, or other),

- area of residence (urban vs. rural),
- referring clinic type (primary care vs. behavioral health)

To allow for formal testing of heterogeneity of the treatment effect across subgroups, a linear model of the AUC summary measures will be fit with indicator variables for the CC treatment arms and adolescent age, the subgroups of interest and the interaction between subgroup and treatment. This study has not been specifically powered to identify differing treatment effects in each of these subgroups. Subgroup analyses will be performed in the ITT analytic population and use imputed outcomes to account for missing data.

6 ENDPOINTS AND COVARIATES

6.1 Primary endpoint

Suicidal behavior is the primary outcome and will be measured using the Harkavy-Asnis Suicide Scale (HASS) Active Suicidal Behavior Subscale. The HASS includes three sub-scales, each of which have been validated for self-report, with strong psychometric properties in adolescents and adults.[16] The Active Suicidal Behavior Subscale of the HASS asks about the frequency of each of five active suicidal behaviors, including suicide attempt planning, and actual and aborted/interrupted suicide attempts. Participants respond using the Likert scale response options used by Asarnow et al. in their validation study,[16] which will be modified to fit the time period based on each survey (past six months will be used for baseline and 12-month outcome surveys, and past 3 months will be used for 3- and 6-month outcome surveys). Each of the sub-scales of the HASS are scored separately. Scoring is completed by summing the number of the associated response category for all the items in the sub-scale. The primary outcome will be the cumulative risk of suicidal behavior assessed using area under the curve of the HASS active suicidal behavior subscale scores over the 12-month

study period. The area under the curve summary measure will be calculated for each participant.

6.2 Secondary endpoints

Secondary outcomes include suicide attempts, suicide ideation, suicide-related ED utilization and hospitalization, and outpatient mental health treatment. All outcomes will be measured at baseline, 3 months, 6 months, and 12 months.

6.2.1 Suicide attempts

Suicide attempts will be measured using the Columbia Suicide Severity Rating Scale (C-SSRS).

6.2.2 Suicidal ideation

Suicidal ideation will be measured using the Passive Suicidal Ideation Subscale of the HASS, which includes 12 items.[16]

6.2.3 Suicide Related ED and hospital utilization

Suicide-related ED utilization and hospitalization will be assessed by self-report and using electronic medical records.

6.2.4 Outpatient mental health treatment

Outpatient mental health treatment will be self-reported and assessed using electronic medical records. Reason for visit/diagnoses may be reviewed in electronic medical records for any inpatient or outpatient encounters that occur during the study period.

6.3 Safety Outcomes and Other Exploratory Endpoints

Exploratory and safety outcomes include thwarted belongingness, perceived burdensomeness, a combined measure of suicidal ideation and behavior, suicide attempts (including actual, aborted, or interrupted attempts), perceived mattering to others, non-suicidal self-injury, all-cause ED utilization and hospitalizations, suicide deaths, and all-cause mortality.

- Thwarted belongingness and perceived burdensomeness will be measured using the 15-item version of the Interpersonal Needs Questionnaire.[19-21]
- A combined indicator of suicidal ideation and behavior will be assessed using the Columbia Suicide Severity Rating Scale self-report 6-item screener for primary care (C-SSRS).[22,23] The C-SSRS self-report screener is widely used in clinical practice, including at SLHS, and will be used to determine eligibility for the study. Participants will also self-complete the C-SSRS at baseline following enrollment, and at 3, 6, and 12 months. The C-SSRS self-report 6-item screener has strong psychometric properties for use with both adolescent and adult and populations, including excellent sensitivity and specificity,[22] convergent validity,[23] and incremental validity.[23]
- Non-suicidal self-injury will be measured with items from the SITBI-R. The General Mattering Scale (GMS)[25-27] will be used to assess the extent to which participants believe they matter to other people.
- All-cause ED utilization and hospitalization will be measured from self-report and electronic medical records.
- The frequency of suicide attempts (including actual attempts, interrupted attempts, and self-aborted attempts) will be measured using the two-item Suicide Attempts Sub-Scale of the HASS.[16]

- The manner and cause/lethal means of suicide deaths and all deaths will be assessed based on electronic medical records and vital records.

6.3.1 Adverse Events and Serious Adverse Events

This protocol does not include tracking of adverse events (AEs) nor real-time tracking of serious adverse events (SAEs) for several reasons. First, the most important events that would be defined as SAEs are expected safety outcomes in the context of study participants experiencing suicidality. All deaths of study participants will be reviewed and assessed to determine whether the cause of death is suicide. We do not anticipate any SAEs beyond those listed as safety outcomes above, but unanticipated SAEs that occur will be reviewed by the DSMB and the IRB. Second, the most important safety question to ascertain in the context of this trial is whether rates of safety outcomes or SAEs are differential across the two intervention groups. Given that the CC2 intervention includes frequent contact with study participants (and the CC1 and UC conditions do not), if the study were to monitor AEs and/or SAEs in real-time, differential rates of ascertainment would be expected. Any attempt to compare rates of AEs or SAEs across intervention groups could be substantially biased due to differential ascertainment. Instead, this protocol will monitor safety outcomes collected at 3, 6, and 12 months as part of routine study outcome assessments through regular DSMB meetings to ensure equal ascertainment of outcomes across intervention groups. This is the most valid and reliable way to review safety data in the context of this pragmatic clinical trial.

6.4 Covariates

We will collect data on age and sex at birth, as well as gender identity and sexual orientation, all of which are strongly associated with risk of suicidal ideation and behavior.[1,28,29] Gender identity and sexual orientation will be assed using questions from CDC's Youth Risk Behavior Surveillance System Questionnaire.[30] Suicide rates differ by

race and ethnicity,[1] which will be self-reported using US Census categories.[31] Religious affiliation and practice is associated with suicidality, but the magnitude and direction of that association differs depending on the religion, and its intersection with socio-cultural factors (e.g.: sexuality).[32] We will include items from a recent Pew Research Religious Landscape Study to assess religiosity.[33] Compared to other adults, active-duty military or veterans face a 57.3% higher adjusted suicide rate.[34] The Health Resources & Services Administration's Office of Rural Health Policy urban-rural designation for census tracts[35] will be used to classify participants as urban or rural residents. Compared to large urban areas, residence in the most rural and remote parts of the US is associated with a 96% higher rate of suicide,[1] lower socioeconomic status,[36] and worse overall health outcomes.[36] Thirty-five of Idaho's 44 counties are classified as rural.[35] Financial crises and low socioeconomic status (SES) are associated with increased suicidality.[50] We will collect data related to SES including income, education, and employment using questions from the US Census Bureau's American Community Survey.[31] Lethal means for planned and actual suicide attempts will be self-reported.[38] Utilization of 988 will be self-reported by participants to determine how different sub-groups & treatment arms used this resource before and during the study.

7 STATISTICAL METHODOLOGY

7.1 Descriptive statistics

Descriptive statistics will be used to summarize survey results including outcomes, baseline demographics, and other exposure variables. This will include number and percentage for categorical variables; for continuous variables mean and standard deviation will be tabulated, with median, first quartile, and third quartile included in at least internal reports.

7.2 Primary analysis

The primary analysis population is the intention to treat population: all randomized participants who complete a baseline survey, grouped by the treatment arm to which they were randomized regardless of treatment delivered. An area under the curve analysis will be used to assess HASS Suicidal Behavior score as a measure of cumulative risk across the 12-month study period. A linear model with indicator variables for the treatment arms and using heteroskedasticity robust sandwich variances will be used in the primary model. Multiple imputation will be used to account for missing outcome data and Rubin's rule will be used to combine results across imputations.[41] The primary analysis will use a gatekeeping procedure: first, the Caring Contact arms will be compared to usual care; if both CC arms are statistically significantly superior to usual care, then a noninferiority comparison will be done for the two Caring Contact study arms using a noninferiority margin of 1.0.

7.2.1 Simulation results

7.3 Secondary analyses

The analyses of secondary outcomes will use linear models similar to the primary analysis; for binary outcomes, a generalized linear model will be fit with the identity link to allow for the estimation of the risk difference. Analyses will be performed using standard statistical software such as R or SAS. A two-sided type-I error of 0.05 will be used as the threshold to determine statistical significance. Confidence intervals will be reported in addition to p-values. All outcomes will be measured at baseline, 3 months, 6 months, and 12 months. The analysis of secondary outcomes will not be adjusted for sequential monitoring since these outcomes will not be formally included in the interim analysis.

7.3.1 Suicide attempts

A count of suicide attempts will be measured using the Columbia Suicide Severity Rating Scale (C-SSRS). The total number of attempts across the 12 month study period will be compared across treatment arms by a negative binomial hurdle model.[38-41] Kaplan-Meier curves will be used to compare the time to first suicide attempt between treatment arms; the log-rank test will be used to test for differences between groups.

7.3.2 Suicide ideation

Suicidal ideation will be measured using the Passive Suicidal Ideation Subscale of the HASS, which includes 12 items.[16] For each participant, the area under the curve will be used to summarize the total burden of suicidal ideation across the 12 month study period. The average AUC will be compared between treatment arms using a linear model with the heteroskedasticity robust sandwich variances.

7.3.3 Suicide-related ED utilization and hospitalization

Suicide-related ED utilization and hospitalization will be assessed by self-report and using electronic medical records. We will present the number and proportion of participants with any suicide-related ED or hospitalization during the course of the 12-month study. The overall risk difference comparing the caring contact arms to usual care will be estimated from a GEE model.

7.3.4 Outpatient mental health treatment

Outpatient mental health treatment will be self-reported and assessed using electronic medical records. Reason for visit/diagnoses may be reviewed in electronic medical records for any inpatient or outpatient encounters that occur during the study period. We will present the number and proportion of participants reporting the use of mental health treatment at each outcome survey time. The risk difference in the proportions

will be estimated at each time point using a GEE model which models all outcomes simultaneously to account for the correlations of any given participant's responses.

7.4 Subgroups

Subgroup analyses will be completed based on

- age at enrollment (12-17; 18-24; 25-49; 50+),
- sex at birth (female vs. male),
- gender identity (cisgender vs. transgender or gender nonconforming),
- sexual orientation (heterosexual vs. homosexual, bisexual, or other), and
- area of residence (urban vs. rural),
- referring clinic type (primary care vs. behavioral health).

To allow for formal testing of heterogeneity of the treatment effect across subgroups, a linear model of the AUC summary measures will be fit with indicator variables for the CC treatment arms, the subgroups of interest, and the interaction between subgroup and treatment. The effect of each CC treatment, compared to usual care, will be reported along with the 95% confidence interval for the difference in means and the p-value testing for the interaction of group with treatment. The subgroup analyses will be performed in the multiple imputed data sets for the ITT population; Rubin's rule will be used to combine results across imputations.[41]

7.5 Exploratory analyses

7.5.1 Thwarted belongingness and perceived burdensomeness

Both thwarted belongingness and perceived burdensomeness will be summarized over time using the AUC summary measure, as in the primary analysis. The difference in means between treatment arms will be estimated from a linear regression model using robust sandwich standard errors.

7.5.2 C-SSRS Suicidal Ideation and Behavior

The AUC will also be used to summarize each participant's C-SSRS suicidal ideation and behavior scores over time. The differences in mean AUC between treatment arms will be estimated from a linear regression using robust sandwich standard errors.

7.5.3 Non-suicidal self-injury

The time to first non-suicidal self-injury will be compared across treatment arms using the Kaplan-Meier method, with the log-rank test used to test for differences between treatment arms.

7.5.4 All-cause ED and hospital utilization

All-cause ED utilization and hospitalization will be measured from self-report and electronic medical records. We will present the number and proportion of participants with any ED or hospitalization utilization during the course of the 12-month study. The overall risk difference comparing the caring contact arms to usual care will be estimated from a GEE model.

7.5.5 Suicide attempts

The occurrence of suicide attempts will be summarized as a binary outcome at each measured timepoint. A GEE model with the identity link function and binomial variance structure will be used to estimate the risk differences between the treatment arms at each time point.

7.5.6 Cause of death

The manner and cause of all deaths, including suicide deaths will be reported in a listing. This will include lethal means for deaths by suicide.

7.6 Safety analysis

The proportion of participants experiencing each of the safety events will be summarized within each treatment arm. We will present the estimated risk difference

and confidence intervals for the risk difference. These analyses will be performed in the safety population.

7.7 HANDLING OF MISSING OUTCOME DATA

Every attempt will be made to minimize the amount of missing data. Regular reports will allow the investigators to assess the amount of missing data as well as the cause of missing or anomalous entries and address these if possible. If missing values cannot be retrieved, the reason for the missingness will be recorded in data comments available through REDCap. Hotline specialists will record the reason for any participant drop-out during the 12 months of follow-up. Missing outcome data due to withdrawal or loss to follow-up will be addressed through multiple imputation. Participants who die by suicide will be imputed to have the highest negative score on a given scale; other participants missing data due to death will have their outcomes imputed.

Hot deck imputation will be used to account for missing 3-, 6-, and 12-month outcomes.[42] Outcomes will be sampled from complete cases on the same treatment arm, baseline survey C-SSRS score (low, moderate, and high risk), and age (12-17; 18-24; 25-49; 50+ (<25, 25+). This method avoids making parametric assumptions about the outcome distribution and preserves the relationship between measures.[42] Twenty complete data sets will be imputed and Rubin's rule will be used to pool the results.[41]

7.8 SENSITIVITY ANALYSES

We will assess the sensitivity of estimates to the imputation approach and variables used in the imputation algorithm, including the possibilities that missing data is more likely to reflect poor outcomes or that the missing data mechanism varies by treatment arm. A tipping point analysis will be included.

The primary analysis will be repeated in the per-protocol population, defined as the subset of participants who were retained for the duration of the study, received only the randomized treatment, and completed the minimum assigned intervention (e.g. received all scheduled text messages for the caring contacts arms).

The HASS active suicide behavior subscale scores at 3, 6, and 12 months will be analyzed using a negative binomial hurdle . This model will include study arm. The baseline HASS active suicidal behavior score will be included as a precision variable.

7.9 PROGRAMMING PLANS

All data cleaning and programming will be done in REDCap, Excel, SAS, R or RStudio.

8 References

1. Ehlman DC, Yard E, Stone DM, Jones CM, Mack KA. Changes in Suicide Rates — United States, 2019 and 2020. *MMWR Morb Mortal Wkly Rep.* 2022;71:306-312. doi:http://dx.doi.org/10.15585/mmwr.mm7108a5

2. Motto JA, Bostrom AG. Follow-Up for Parasuicidal Patients. *Psychiatric Services* 2001;52(9):1254-1254. doi:10.1176/appi.ps.52.9.1254-a

3. Skopp NA, Smolenski DJ, Bush NE, et al. Caring Contacts for Suicide Prevention: A Systematic Review and Meta-Analysis. *Psychological services* 2023;20(1):74-83. doi:10.1037/ser0000645

4. Radin AK. Suicide Prevention Among Recipients of Care (SPARC). ClinicalTrials.gov Web site. https://clinicaltrials.gov/ct2/show/NCT04893447. Updated 2021. Accessed July 21, 2022

5. Radin AK. Mental Health Among Patients, Providers, and Staff During the COVID-19 Era (MHAPPS). ClinicalTrials.gov Web site. https://clinicaltrials.gov/ct2/show/NCT04700137. Updated 2022. Accessed July 21, 2022

6. Radin, AK., Shaw, J, Brown, SP, Torres, J, Harper, M, Flint, H, Fouts, T, McCue E, Skeie, A, Peña, C, Youell, J, Doty-Jones, A, Wilson, J, Flinn, L, Austin, G, Chan, KCG, Zheng, Z, Fruhbauerova, M, Walton, M, Kerbrat, A, Comtois, KA. Comparative Effectiveness of Two Versions of a Caring Contacts Intervention in Healthcare Providers, Staff, and Patients for Reducing Loneliness and Mental Distress: A Randomized Controlled Trial. *J Affect Disord.* 2023; 331:442-451. doi: 10.1016/j.jad.2023.03.029.

7. Hassanian-Moghaddam H, Sarjami S, Kolahi A, Carter GL. Postcards in Persia: Randomised Controlled Trial to Reduce Suicidal Behaviours 12 Months After Hospital-Treated Self-Poisoning. *British Journal of Psychiatry* 2011;198(4):309-316. doi:10.1192/bjp.bp.109.067199

8. Carter GL, Clover K, Whyte IM, Dawson AH, Este CD. Postcards from the EDge Project: Randomised Controlled Trial of an Intervention Using Postcards to Reduce Repetition of Hospital Treated Deliberate Self Poisoning. *BMJ.* 2005;331(7520):805-807. doi:10.1136/bmj.38579.455266.E0

9. Comtois KA, Kerbrat AH, DeCou CR, et al. Effect of Augmenting Standard Care for Military Personnel with Brief Caring Text Messages for Suicide Prevention: A Randomized Clinical Trial. *JAMA Psychiatry*. 2019;76(5):474-483. doi:10.1001/jamapsychiatry.2018.4530

10. Kittelson JM, Emerson SS. A unifying family of group sequential test designs. *Biometrics* 1999;55:874-82.

11. Pampallona S, Tsiatis AA. Group sequential designs for one-sided and two-sided hypothesis testing with provision for early stopping in favor of the null hypothesis. *J Stat Planning and Inference* 1994;42:19-35.

12. Whitehead J. On the bias of maximum likelihood estimation following a sequential test. *Biometrika* 1986;73:573-81.

13. Emerson S, Fleming T. Parameter estimation following group sequetial hypothesis testing. *Biometrika* 1990;77:875-92.

14. Burington BE, Emerson SS. Flexible Implementations of Group Sequential Stopping Rules Using Constrained Boundaries. *Biometrics* 2003;59(4): 770-777. https://www.jstor.org/stable/3695315

15. Mosio I. Mosio - Text Messaging Solution for Clinical Trials. https://www.mosio.com/. Accessed November 30, 2022

16. Asarnow J, McArthur D, Hughes J, Barbery V, Berk M. Suicide Attempt Risk in Youths: Utility of the Harkavy-Asnis Suicide Scale for Monitoring Risk Levels. *Suicide & Life-threatening Behavior* 2012;42(6):684-698. doi:10.1111/j.1943-278X.2012.00122.x

17. Fox KR, Harris JA, Wang SB, Millner AJ, Deming CA, Nock MK. Self-Injurious Thoughts and Behaviors Interview-Revised: Development, reliability, and validity. *Psychol Assess.* 2020;32(7):677-689. doi:10.1037/pas0000819

18. Gratch I, Fernandes SN, Bell K, et al. Self-Injurious Thoughts and Behaviors Interview-Revised (SITBIR): Reliability, Validity, and Inter-Informant Agreement in an Adolescent Sample. *Journal of Clinical Child and Adolescent Psychology*. 2022;51(4):484-494. doi:10.1080/15374416.2021.190122915.

19. Hill RM, Rey Y, Marin CE, Sharp C, Green KL, Pettit JW. Evaluating the Interpersonal Needs Questionnaire: Comparison of the Reliability, Factor Structure, and

Predictive Validity across Five Versions. *Suicide and Life-Threatening Behavior* 2015;45(3):302-314. doi:10.1111/sltb.12129

20. Lutz J, Fiske A. Perceived Burdensomeness in Older and Younger Adults: Evaluation of the Psychometric Properties of the Interpersonal Needs Questionnaire. *Journal of Clinical Psychology* 2017;73(9):1179-1195. doi:10.1002/jclp.22415

21. Van Orden KA, Cukrowicz KC, Witte TK, Joiner TE. Thwarted Belongingness and Perceived Burdensomeness: Construct Validity and Psychometric Properties of the Interpersonal Needs Questionnaire. *Psychol Assess.* 2012;24(1):197-215. doi:10.1037/a0025358

22. The Columbia Lighthouse Project / Center for Suicide Risk Assessment. The Columbia Suicide Severity Rating Scale (C-SSRS) Supporting Evidence. Columbia Lighthouse Project Website Web site. https://cssrs.columbia.edu/wp-content/uploads/CSSRS_Supporting-Evidence_Book_2018-10-10.pdf. Updated 2018. Accessed November 30, 2022

23. Viguera AC, Milano N, Laurel R, et al. Comparison of Electronic Screening for Suicidal Risk with the Patient Health Questionnaire Item 9 and the Columbia Suicide Severity Rating Scale in an Outpatient Psychiatric Clinic. *Psychosomatics (Washington, D.C.)* 2015;56(5):460-469. doi:10.1016/j.psym.2015.04.005

24. Posner K, Brown GK, Stanley B, et al. The Columbia–Suicide Severity Rating Scale: Initial Validity and Internal Consistency Findings From Three Multisite Studies With Adolescents and Adults. *American Journal of Psychiatry* 2011;168(12):1266-1277. doi:10.1176/appi.ajp.2011.10111704

25. Rosenberg M, McCullough B. Mattering: Inferred Significance and Mental Health among Adolescents. *Research in Community and Mental Health* 1981;2:163-182. https://search.proquest.com/docview/61082562

26. Davis SM, Lepore SJ, Dumenci L. Psychometric properties and correlates of a brief scale measuring the psychological construct mattering to others in a sample of women recovering from breast cancer. *Qual Life Res.* 2019;28(6):1605-1614. doi:10.1007/s11136-019-02118-8

27. DeForge BR, Barclay DM. The Internal Reliability of a General Mattering Scale in Homeless Men. *Psychological Reports* 1997;80(2):429-430. doi:10.2466/pr0.1997.80.2.429

28. Johns MM, Lowry R, Andrzejewski J, et al. Transgender Identity and Experiences of Violence Victimization, Substance Use, Suicide Risk, and Sexual Risk Behaviors Among High School Students — 19 States and Large Urban School Districts, 2017. *MMWR. Morbidity and mortality weekly report.* 2019;68(3):67-71. doi:10.15585/mmwr.mm6803a3

29. Johns MM, Lowry R, Haderxhanaj LT, et al. Trends in Violence Victimization and Suicide Risk by Sexual Identity Among High School Students — Youth Risk Behavior Survey, United States, 2015–2019. *Morbidity and mortality weekly report. Supplement.* 2020;69(1):19-27. doi:10.15585/mmwr.su6901a3

30. Centers for Disease Control and Prevention, US Department of Health and Human Services. 2023 State and Local Youth Risk Behavior Survey. https://www.cdc.gov/healthyyouth/data/yrbs/pdf/2023/2023_YRBS_Standard_HS_Questionnaire.pdf. Updated 2022. Accessed November 30, 2022

31. United States Census Bureau. The American Community Survey. https://www2.census.gov/programs-surveys/acs/methodology/questionnaires/2022/quest22.pdf. Updated 2022. Accessed November 30, 2022

32. Wu A, Wang J, Jia C. Religion and Completed Suicide: a Meta-Analysis. *PLoS ONE* 2015;10(6):e0131715. doi:10.1371/journal.pone.0131715

33. Pew Research Center. Pew Research Center 2014 Religious Landscape Study (RLS-II) Topline. Pew Research Center Web site. https://www.pewresearch.org/religion/wpcontent/uploads/sites/7/2015/05/RLS-II-FINAL-TOPLINE-FOR-FIRST-RELEASE.pdf. Updated 2014. Accessed November 30, 2022

34. U.S. Department of Veterans Affairs, Office of Mental Health and Suicide Prevention. 2022 National Veteran Suicide Prevention Annual Report. U.S. Department of Veterans Affairs. 2022

35. US Department of Agriculture Economic Research Service. Rural-Urban Commuting Area Codes.

36. Centers for Disease Control & Prevention, US Department of Health and Human Servicees. CDC Health Disparities & Inequalities Report. *Morbidity & Mortality weekly Report (MMWR).* 2013;62(3)

37. Atkins DC, Baldwin SA, Zheng C, Gallop RJ, Neighbors C. A Tutorial on Count Regression and Zero-Altered Count Models for Longitudinal Substance Use Data. Psychology of addictive behaviors. 2013;27(1):166–177. doi:10.1037/a0029508

38. Noh M, Lee Y. Extended negative binomial hurdle models. Statistical methods in medical research. 2019;28(5):1540–1551. doi:10.1177/0962280218766567

39. Feng CX. A comparison of zero-inflated and hurdle models for modeling zero-inflated count data. J Stat Distrib App. 2021;8(1):1–8. doi:10.1186/s40488-021-00121-4

40. Saffari SE, Allen JC. Bivariate negative binomial regression model with excess zeros and right censoring: an application to Indonesian data. Journal of applied statistics. 2020;47(10):1901–1914. doi:10.1080/02664763.2019.1695761

41. Rubin DB. *Multiple Imputation for Nonresponse in Surveys.* John Wiley & Sons, Inc; 1987. DOI:10.1002/9780470316696

43. Andridge RR, Little RJA. A Review of Hot Deck Imputation for Survey Non-response. *International Statistical Review* 2010;78(1):40-64. https://doi.org/10.1111/j.1751-5823.2010.00103.x