Study Protocol and Statistical Analysis Plan

Date: 02/04/2025

National Clinical Trials #: NCT06699251

Study Title: Effectiveness of Brief Intervention in Primary Care for Diverse Young People (Chat)

INVESTIGATOR/AUTHOR PAGE

Co-Principal Investigator	Elizabeth D'Amico, Ph.D.
Title/Position	Senior Behavioral Scientist
Organization	RAND
Location	Santa Monica, CA
Co-Principal Investigator	Alina Palimaru, Ph.D.
Title/Position	Policy Researcher, Full
Organization	RAND
Location	Santa Monica, CA
Author	Michael Woodward
Title/Position	Research Project Manager
Organization	RAND
Location	Santa Monica, CA
Version	1.0

Abbreviations

AC	Alcohol and cannabis
ANOVA	Analysis of variance
AOD	Alcohol and other drug
AUDIT	Alcohol Use Disorders Identification Test
BI	Brief intervention
CRAFFT	Car, Relax, Alone, Forget, Friends/Family, Trouble Questionnaire
CUD	Cannabis use disorder
CUDIT-R	Cannabis Use Disorders Identification Test – Revised
EUC	Enhanced usual care
FDR	False discovery rate
HTE	Heterogeneous treatment effect
KWIC	Keyword in context
LGM	Latent growth modeling
MLR	Maximum likelihood with robust standard errors
NNT	Number needed to treat
PC	Primary care
RCT	Randomized controlled trial
SGD	Sexual and Gender Diverse
WLSMV	Weighted least squares with means and variances adjusted

Study Protocol

A clinical study or research protocol guides the study and associated data collection and analysis in a productive and standardized manner and is carefully designed to safeguard the participants' health and answer specific research questions. The protocol should describe the following aspects of the study.

1. What the study will do

Our patient-centered aims will provide critical real-world information for providers and policy makers on how to best implement screening and brief intervention (BI) in primary care (PC) with the overall goal of addressing barriers to improve health equity for populations who are medically underserved. We will leverage our established partnerships with two healthcare organizations to address the following aims.

Aim 1. Elicit stakeholder perspectives via qualitative interviews with teens (N=60), providers (N=10), and parents (N=20) to inform:

a) how racial and ethnic and sexual and gender diverse (SGD) experiences vary for young people, and how that may affect AC use behaviors, which can help enhance acceptability by tailoring Chat to patient and stakeholder values and preferences (e.g., including discussion with teens about how discrimination may affect their AC use),

b) feasibility and potential sustainability of Chat in PC settings, and

c) adaptations of survey scales to capture dimensions relevant to stakeholders (e.g., social determinants of health).

Aim 2a. Compare effectiveness of Chat to Enhanced Usual Care (EUC) in decreasing alcohol and cannabis (AC) use, AC problems, and time spent around peers who use AC, and increasing motivation to change, resistance self-efficacy, and health service use over a one-year period for teens.

Aim 2b. Examine heterogeneity of treatment effects by race and ethnicity, SGD identity, and substance use severity.

Aim 3. Assess patient experience and satisfaction, specifically which components of Chat teens valued most, any differences across patient groups, and how these may relate to their outcomes.

Aim 4. Assess provider and teen perspectives on their experiences, barriers and facilitators, and recommendations regarding the implementation, effectiveness, scalability, and sustainability of Chat.

2. How it will be done

Young people from the PC clinics who meet eligibility criteria, screen in at-risk, and have parental consent and also assent to participate in the study will be randomly assigned to either Chat or EUC after their baseline survey, and followed up at 3-, 6-, and 12 months (Aim 2). Our design is patient-centered, with significant opportunities for patient and stakeholder data elicitation to inform intervention refinement towards acceptability, feasibility, and sustainability (Aim 1); assess patient experience and satisfaction early during the intervention (Aim 3); and ascertain teen and provider post-intervention perspectives about sustainable implementation in PC settings (Aim 4).

3. Why it is being done

Identifying and intervening with teens who use AC regularly or experience problems from AC use is a significant public health priority given increased likelihood of psychosocial, health, financial, and emotional problems that can occur later in life for those who continue to use. Work in this area has shown that AC use during this key developmental period is associated with numerous risk behaviors, including unsafe driving, condomless sex, and a greater number of co-occurring problems across physical, psychological and relational domains. Furthermore, AC use during this period can affect subsequent educational and employment opportunities, which can have lasting effects into adulthood.

There is great need to address AC use among adolescents, particularly those who are racially and ethnically diverse and those who report SGD identity. Given that most teens see a PC provider annually, the PC setting offers an important opportunity to reach many teens that otherwise would not be reached and address AC use through a motivational BI.

4. How many people will be in the study

We expect to enroll 500 young people in the randomized controlled trial.

5. Who is eligible to take part in it

Teens aged 12-17 will be eligible to participate if they screen in as at-risk using the CRAFFT.

6. What study therapy or other interventions will be given

None.

7. What tests will be done and how often

No tests will be done. Survey data will be collected at baseline, 3 months, 6 months and 12 months.

8. What information will be collected

From baseline to 3 months, to 6 months, and to 12 months, we will measure change in:

a. Frequency of alcohol use.

b. Frequency of cannabis use.

c. Consequences of alcohol use.

d. Consequences of cannabis use.

e. Time spent around peers who use alcohol.

f. Time spent around peers who use cannabis.

g. Motivation to change substance use.

h. Resistance self-efficacy.

We will also measure use of behavioral health care services.

Statistical Analysis Plan (SAP)

The statistical analysis plan outlines a proposed method to describe and define the following aspects of the study.

1. Statistical aspects of the clinical trial design

Aim 2a. We will compare effectiveness of teens who receive Chat (n-250) to those who receive EUC (n=250) in decreasing AC use, AC consequences, time spent around peers who use AC, and increasing motivation to change, resistance self-efficacy, and health service use over a one-year period.

Aim 2b. We will examine heterogeneity of treatment effects across subgroups, i.e., race and ethnicity, sexual and gender diverse identity, and substance use severity.

Aim 3. We will assess patient experience and satisfaction, specifically which components of Chat teens valued most, any differences across patient groups, and how these may relate to their AC outcomes.

Aim 4. We will assess provider and teen perspectives on their experiences, barriers and facilitators, and recommendations regarding the implementation, effectiveness, scalability, and sustainability of Chat.

2. Process of data selection for analyses

Data will be selected for analysis to address study aims focused on evaluating primary and secondary outcomes.

3. Detailed analyses of data items

Descriptives. As a first step, we will examine descriptive statistics and frequencies for evidence of sparseness for categorical data and for non-normality (using plots, examination of skewness, kurtosis, etc.) for continuous variables. Where sparseness exists in categorical variables, we will collapse as necessary to produce cell sizes sufficient for analysis. Where non-normality is evident, variables may be transformed. Outliers may be recoded or omitted if necessary. Given that participant eligibility/inclusion depends on screening positive for alcohol and/or cannabis use, we do not anticipate a preponderance of 0's for outcomes such as use and consequences. However, upon inspection of data, should we find an excess of 0's, we are prepared to handle this using two-part growth models, which can be easily estimated in Mplus.

Baseline equivalence across experimental groups. We will evaluate comparability of experimental groups with respect to potential confounders. We will use categorical methods of analysis (e.g., cross-tabulations, chi-square) to compare groups for discrete data (e.g., employment, school status). We will use analysis of variance (ANOVA) or t-tests to test for homogeneity of groups for continuous data at baseline. If a statistically significant difference is found, we will include covariates in all subsequent analyses. We will address missing data using multiple imputation and/or full information maximum likelihood estimation.

4. Procedures and methods employed for analyzing the data

Aim 2a. Compare effectiveness of Chat to EUC in decreasing AC use, AC problems, and time spent around peers who use AC, and increasing motivation to change, resistance self-efficacy, and health service use over a one-year period for teens.

We will use intention to treat analyses and will attempt to follow up with all individuals regardless of Chat or EUC completion. We will test longitudinal change and comparisons between groups using SAS Proc Mixed for continuous outcomes and SAS Proc Glimmix/NLMixed for categorical outcomes. Alternatively, we will examine change using multiple-group latent growth models or known groups growth mixture models in Mplus using robust maximum likelihood (MLR) or weighted least squares with means and variances adjusted (WLSMV) estimation to accommodate both continuous and categorical outcomes. Consistent with most randomized controlled trials (RCTs), we will examine each outcome separately within one of the proposed frameworks.

Data Missingness. Regarding minimizing data missingness, we have various ways we minimize data missingness. For individual items (i.e., outcomes) within a survey that are unanswered, we program a prompt that notifies the participant an answer is missing. It asks participants to answer the question if it was left blank unintentionally but allows them to move forward without answering if their intention was to refuse to answer. At the participant level, when participants do not respond to a survey completion request, we follow-up with them various times in various modes to encourage participation.

Sample size and power: We conservatively compute estimated power based on the final projected sample size accounting for attrition at 12-month follow-up. Based on our previous work following young people over time, we estimate 80% retention at 12-month follow-up – a final sample of 400 participants (n=200 per arm). Additionally, we consider prior reported standardized effects sizes for Chat (e.g., 0.21 to 0.86 for continuous measures of cannabis use, number of friends who use, intentions, and consequences) in the estimation of power. With these minimum sample sizes per arm, assuming a correlation between repeated assessments of 0.50, four timepoints, and alpha of .05, we have 80% power to detect a standardized effect size ($\underline{d}$) of 0.222 between groups and standardized effects size (d) of 0.116 within groups; thus, we are powered to detect very small effects using conventional standards for Cohen's d. We will also compute and report false discovery rates using the Benjamini-Hochberg false discovery rate (FDR) procedure. In the Benjamini-Hochberg FDR procedure, individual p-values obtained from all comparisons/tests (significant and non) are placed in ascending order and assigned a rank (1 - # of tests conducted). Then Benjamini-Hochberg critical values are computed using the formula $(i / m)*Q$ where i is the item rank, m is the number of total tests conducted and Q is the desired p-value (in this case 0.05). Then each individual obtained p-value is compared to the Benjamini-Hochberg critical value. Thus, each individual p-value is evaluated against a corresponding adjusted p-value. As such, the current power calculations remain unchanged as the correction for multiple tests occurs on the obtained p-values rather than adjusting the desired p-value for power calculations. That said, should the overly stringent Bonferroni approach be used wherein alpha is adjusted based on the 4 primary outcomes, the corrected alpha used in power calculations would be .01. If we used a Bonferroni approach, the effect on study power would be minimal as detectable between groups effects would increase from .222 to .272 and for within group effects from .111 to .1393. We present these new calculations here in good faith; however, we will use the FDR approach to adjust for multiple comparisons for all proposed analyses.

Aim 2b. Examine heterogeneity of treatment effects by race and ethnicity, SGD identity, and substance use severity.

We will also explore whether treatment effect estimates are heterogeneous across subgroups (i.e., SGD identity, race and ethnicity, and substance use severity) based on the baseline survey. Heterogeneous treatment effect (HTE) analyses will be conducted by modifying Aim 2a analysis models to include as right-hand side variables treatment assignment, subgroup variable and their interaction. Significance of the interaction term will indicate subgroup differences in treatment effects. We selected these participant subgroups based on our own and other rigorous prior research that demonstrates significant differences in terms of intervention outcomes for young people at risk for AC use based on marginalized group status (e.g., SGD, race and ethnicity) and substance use severity.

Measures for HTE analyses. The Alcohol Use Disorders Identification Test (AUDIT) focuses on frequency and consequences of drinking. We will use the modified version that is developmentally appropriate for young people. The AUDIT has 10 items (e.g., have you or someone else been injured as a result of your drinking?) and a cut score of 4 or higher with this age group has the best specificity and sensitivity for problem use. The Cannabis Use Disorders Identification Test – Revised (CUDIT-R) has 8 items (e.g., how many hours were you 'stoned' on a typical day when you were using cannabis?) and evaluates problem severity by distinguishing between different levels of cannabis use, cannabis use disorders (CUD), and stage of change. A score of 8 or higher indicates hazardous cannabis use, and scores of 12 or greater indicate a possible CUD.

Sample size and power. We focus our power calculation for pre-specified HTE analyses for which we estimate we will have 200 participants per condition at 12 months, after accounting for 80% retention. For all pre-specified HTE tests, we use an alpha level of 0.05 (two-tailed) and determine detectable effect sizes for the interaction with 80% power. We have 80% power to detect an interaction with a small effect (Cohen's d = 0.292; OR = 1.69) in a 2x2 design (e.g., treatment x SGD identity); small effect (Cohen's d = 0.323; OR = 1.80) in a 2 x 3 design (e.g., treatment x use severity [low, medium, high]); and a small effect (Cohen's d = 0.362; OR = 1.93) in a 2 x 5 design (e.g., treatment x race and ethnicity). To further account for multiple statistical testing with secondary HTE analyses and the aforementioned Aim 2a analyses, we will also compute and report false discovery rates using the Benjamini-Hochberg FDR procedure.

Aim 3. Assess patient experience and satisfaction, specifically which components of Chat teens valued most, any differences across patient groups, and how these may relate to their AC outcomes. Mixed methods data collection.

To ascertain the level and type of satisfaction and experience that teens have with the Chat intervention, we will employ a convergent mixed-methods approach, using written narratives and survey data collected at 3-month follow-up. Teens will answer closed-ended questions about satisfaction with quality and delivery of the interventions, outcomes, and perceived patient-centeredness, followed by open-ended questions, inviting narrative feedback on satisfaction and other dimensions of experience.

Mixed methods analysis. We will analyze open-ended comments for themes relating to satisfaction and experience and will compare themes with correspondent quantitative data using joint displays of themes and survey ratings. We will also conduct textual analysis of open-ended narratives about satisfaction and experience using the Keyword in Context (KWIC) approach in NVivo. NVivo will produce a raw frequency of distinctive words used by each participant to explain why they are or are not satisfied with the Chat intervention. Through the KWIC function, the most frequently used distinctive words will be selected and analyzed in context – that is, 15 words preceding and 15 words following each frequently used word. We will produce a list of the top five most distinctive words from narratives classified as "satisfied" and the five most distinctive words from those classified as "not satisfied." The KWIC analysis will help us contextualize

frequently used words and understand valence and relationship to satisfaction rating. Data collected and analyzed in Aim 3 are crucial as they can help ensure that later dissemination and implementation efforts take advantage of the most useful components of the intervention and can make adjustments to best accommodate teens' needs and interests.

Aim 4. Assess provider and patient perspectives post-RCT. Interviews with teens and providers.

This explanatory approach will occur at the end of the RCT, with a view to understanding provider and teen experiences, barriers and facilitators, and recommendations regarding the implementation, effectiveness, scalability, and sustainability of Chat. We will recruit a subsample of 20 teens to participate in remote (phone or Zoom) interviews. These will be teens who participated in Chat stratified by location (Pittsburgh and Los Angeles) and outcomes (by no improvement in outcomes and outcome improvement). Specifically, teens will be asked about perceived effects of the intervention and any observed changes in their behavior, as well as suggestions for future implementation. We will also recruit up to 10 providers from AltaMed and UPMC to elicit reflections on their role in the study, the extent to which they think Chat was successfully implemented within existing workflows at their clinics, lessons learned about how to overcome any perceived barriers in the future, and any other factors relating to Chat's scalability, and sustainability. Each participant will be offered a gift card incentive, which will be emailed or mailed to them shortly after the interview is completed.

5. Planned presentation of results in formats such as tables, listings, and figures

Results will be depicted in tables or figures as appropriate.

6. Plans for interim and final analysis and statistical analysis of the primary and secondary variables and other data

All analyses detailed above in Sections 3 and 4 including timepoint specific follow-up as well as overall longitudinal models will compare intervention conditions on change in primary outcomes. As follow-up data are collected, interim analysis of intervention effects will be evaluated.