Supplementary Online Content


eAppendix 1. Study Design of the IIDEA RCT

eAppendix 2. Description of the Intervention and Fidelity

eAppendix 3. Analytical Methods—Data Preparation

eAppendix 4. Analytical Methods—Data Analysis

eFigure 1. IIDEA Study Flow

eFigure 2. Heterogenous Treatment Effects on Substance Misuse Outcomes Across Different Subgroups

eTable 1. List of Assessment Measures/Variables

eTable 2. Analysis of Participation Rates (Receiving 6+ Sessions) in IIDEA Intervention on Main Outcomes at 6-Month Follow-up

eTable 3. Intent-to-Treat Analysis of IIDEA Intervention on Main Outcomes by Site

eTable 4. Intent-to-Treat Analysis of IIDEA Intervention on Main Outcomes by Received Intervention Type

eTable 5. Statistically Significant Baseline Differences Between Participants With Missing/Incomplete Assessments and Participants With Complete Assessments

This supplementary material has been provided by the authors to give readers additional information about their work.
STUDY DESIGN SUMMARY
We conducted the IIDEA randomized clinical trial between September 2014 and February 2017, with mostly Latino immigrants in Massachusetts and at two sites in Spain. Based on the promising results of the AC-OK as a screener for dual diagnosis in this population, we used the AC-OK as the screening tool for the clinical trial. Eligible participants needed to endorse two mental health items and two substance use items on the AC-OK, and not be in specialty behavioral healthcare in the past 3 months or have an appointment planned. The trial design included five research assessments: baseline, and follow up assessment at two months, four-months, six-months, and twelve-months from baseline. Participants were also administered a Computerized Adaptive Testing interview for mental health (CAT-MH), funded by a research supplement (NIMH; PI: Gibbons). Participants were randomized into either a 10-session intervention or a control group that received regular check-ins with a care manager. Participants were also provided access to HIV and STI testing by the study team and urine testing for drug metabolites as part of administrative supplement to the NIDA grant (PI: Alegria).

STUDY SETTING
Participants for the randomized controlled trial were identified through recruitment in primary care clinics, community sites, and emergency departments, as well as from referrals from patients to others in their social network who might benefit from the program. Recruitment was conducted between September 2014 and May 2016 in Boston, Madrid and Barcelona. Approval was obtained from the institutional review boards of participating institutions. We held a series of presentations with Directors and staff in the clinics and community sites to introduce the study and determine optimal procedures for outreach and recruitment of patients.

SAMPLING, RECRUITMENT AND CONSENT
Research Assistants (RAs) approached potential participants in person in the waiting room of clinics and community agencies. In certain sites, RAs took contact information for patients and followed up by phone to either screen or administer informed consent, depending on the permission granted by each institutional review board. In some cases, RAs called a list of patients made available by clinical referral, directors of the site, or by access to electronic data.

Research staff first obtained informed consent from participants and then utilized the study screening to identify patients who were eligible for the study using the AC-OK short screener. We asked participants’ permission to contact them through either their PCP or two close contacts that had had the most frequent or stable contact with the participant over the past six months. We collected an address to send intervention materials and follow up letters and establish contact with the participant.

Patients who were NOT eligible received only the short screener and the CAT-MH interview and were compensated $20. The CAT-MHTM is a brief computerized adaptive mental health screening that was validated in Spanish with this sample as part of a NIMH research supplement.

Patients who were eligible received the short screener and were scheduled for a follow up baseline interview. The baseline interview included instruments designed to identify mental health problems, substance misuse and HIV risk behaviors, as well as socio-demographics, cultural, contextual and social factors, medication use, chronic conditions, disability in daily activities, health literacy, language proficiency, access to health services, migration history, and assessments of illness management/recovery and mindfulness. At baseline, they were also administered the CAT-MH interview and a urine test for drug metabolites to determine drug use. They were administered a capacity to consent form to ensure they understood and could take part in the full trial. They were compensated $40-50 for the time spent answering the assessment, depending on what assessment they were completing (see below). Upon completion of baseline, participants were randomized into either the intervention or control condition and began the trial. Additional interviews were administered by research assistants blind to study condition at 2 months following baseline, 4 months, 6 months, and 12 months. At each interview, an urine test was administered for drug metabolites.

Emergency protocol: We used an emergency protocol throughout the screening phase and in the trial. If participants responded affirmatively to questions in the screener or interview related to suicidal thoughts, they were administered the Paykel suicidality screener after the interview. The emergency procedure was prompted if a patient endorsed a 4
or 5 on the Paykel suicidality screener over the past 30-days. The questions include: “Has there been a time in the last 30 days when you reached the point where you seriously considered taking your own life, or perhaps made plans how you should go about doing it?” = yes or “Has there been a time in the past 30 days/since the last interview when you made an attempt on your own life?” = yes. In Massachusetts, study staff connected the patient by phone or in person to the BEST emergency services team, which performs an assessment of patient safety and provides referrals in the event immediate care is needed. In the two sites in Spain, participants were referred to clinical staff overseeing the study, who assessed and helped connect with emergency care if needed. Participants with active suicidality could be contacted after 30 days for reassessment to see if they could safely participate in the trial, if they were not actively receiving psychotherapy.

Certificate of Confidentiality: We received a Certificate of Confidentiality from NIDA for the US-based participant in this project.

RESEARCH PROCEDURES
We recruited bilingual and bicultural study staff who would be engaging, non-judgmental, persistent, and not easily discouraged. Given the sensitivity around substance use and mental health issues and potential concerns about immigration status and discrimination, we selected bilingual Spanish/English staff, who were also Latino and were sometimes born outside of the US or Spain, to build trust and familiarity with participants. All RAs received a comprehensive training to ensure they could conduct the interview effectively and make the participant feel comfortable. RAs followed 4 steps as part of their training 1) review of assessment questionnaire and materials, 2) remote training via WebEx, 3) conduct of at least two successful role plays of the interviews; and 4) intense quality control review of the audio recording of first 2 patient interviews, including receiving detailed feedback from a study supervisor. During the study, all interviews were recorded and periodically evaluated under quality control procedures. A minimum of 15% of all the interviews in the study were randomly selected, distributed to the quality control team and evaluated by the project coordinator of each site for feedback to the interviewer, and corrections if necessary.
eAppendix 2. Description of the IIDEA Intervention and Fidelity

The Integrated Intervention for Dual Problems and Early Action (IIDEA) is a 10-12 session, transdiagnostic, manualized therapy. It was developed to address elevated symptoms of mental health (depression, anxiety, trauma) and substance use (drug, alcohol, and benzodiazepine misuse) in Latino adults.

IIDEA uses a transdiagnostic model, based in research showing that Latino immigrants have a range of mild to moderate co-occurring behavioral symptoms, but low rates of mental health services use and treatment engagement\textsuperscript{4-8}. The intervention employs evidence-based components shown effective in transdiagnostic applications (such as Cognitive Behavioral Therapy; CBT). The intervention seeks to improve engagement among Latinos by addressing cultural and psychosocial barriers through offering sessions in the language of the participant, either in-person or by phone, and scheduling sessions at the convenience of the patient. Additionally, it incorporates mindfulness practice to improve awareness of thoughts and reactions to stress, together with CBT and strategies for emotional regulation\textsuperscript{9-12}. Finally, it includes components for HIV/STI prevention by offering HIV/STI counseling and testing and discussing common myths about these illnesses.

The intervention incorporates information from previous manuals that have demonstrated efficacy in different components. For instance, along with co-authors Fortuna and Mueser, we adapted and expanded components of Mindfulness Based Cognitive Therapy for the Treatment of PTSD and SUD in Adolescents (MBCT-DUAL)\textsuperscript{13} and CBT for PTSD\textsuperscript{14}. These manuals demonstrated efficacy in improving PTSD, anxiety, depression symptoms and substance use severity\textsuperscript{13,15}. To address HIV prevention, we selected components of two interventions proven to be effective in decreasing HIV risk: CONNECT\textsuperscript{16-18} and RESPECT\textsuperscript{19-21}, and motivation interviewing techniques were incorporated throughout the intervention\textsuperscript{22}. The IIDEA program includes a participant’s manual containing information presented during the session and exercises to practice during the week.

The program is structured in 10-12 sessions that are expected to take one hour per session, administrated weekly within a maximum period of 6 months. The first session emphasizes engaging the participant through motivational interviewing, cultural formulation and assessment of barriers to care. It also provides an overview of the program and psychoeducation on the identified symptoms. In the second session, the clinician introduces the concept of mindfulness and its practice as well as the principles of CBT, including addressing patterns of negative thoughts, and their relationship with feelings and actions. In the third session, the 5 steps of cognitive restructuring are presented and practiced, and this work continues throughout the 3 subsequent sessions and suplemental activities. Fourth, fifth and sixth sessions emphasize drug use psychoeducation, relapse prevention techniques and effective communication. Sessions seven and eight work on HIV and STI risk behavior prevention, including pre- and post-test counseling and testing for HIV, chlamydia and gonorrhea. Finally, the ninth and tenth session focus on reviewing techniques and building a self-care plan. Clinicians also offer referrals for continued therapy for patients who need it or request ongoing care.

The intervention is delivered in-person or by phone using participant’s preferred language (English or Spanish). In cases where the participants have childcare or illness constraints, home visits are offered.

Through an iterative process, a fidelity rating instrument was developed to include items common across all sessions and specific to the skill sets for each session. The instrument consisted of a quantitative section that measured the completion of the session components and a qualitative section that rated the quality of session delivery by the clinician. The first section scored each component with 2 (indicating that the clinician addressed the component fully), 1 (indicating partial fulfilment), or 0 (indicating that the clinician did not address the component at all). The qualitative section of the fidelity scale included four questions concerning the quality of therapy administration, which applied to all sessions. These covered: Patient understanding, clinician empathy with the patient, time management and overall ability to deliver the key components of the session. These items were scored out of 3, and were given the following rating instructions: 0 = Not at all (the clinician does not carry out any of the skills required; the clinician does not cover any of the important themes of the session); 1= Poor (the patient does not understand most of the important themes); 2= Acceptable (the patient does not understand all of the important themes); 3 = Good (The clinician fulfills all of the requirements. Although there could be some global criticisms, all the elements were reasonably well covered, and the patient understands and can apply the contents).


**eAppendix 3. Analytical methods – data preparation**

Some variables had missing data due to drop-outs, missing assessments, lack of response to questions, or participant assessment of a question as not applicable. To account for missing data, we used multiple imputation methods in Stata version 14.2 via the *mi impute chained* command. The multiple imputations were carried out in three steps: first, we imputed the missing data for all the variable considered, creating 20 imputed datasets that each consist of 341 participants with four follow-up assessments per participant. Second, we ran analysis on each individually imputed data set; third, we aggregated the individual estimate to obtain final estimates and adjusted standard errors for the uncertainty due to imputation.

Each imputation was carried out using the chained equations method. In what follows, we refer to variables with missing data as incomplete variables and those with non-missing data as complete variables. Each incomplete variable $x_j$ was specified as a conditional function ($g_j$) given the set of all other variables used in the imputation, comprising both incomplete ($x_1,...,x_m$) and complete variables ($Z$). We fitted the conditional model $g_j$ to generate the predicted values of $x_j$ using an iterative method. Specifically, each incomplete variable $x_j$ was iteratively estimated and in each iteration the variable $x_j^t$ was then updated to $x_j^{t+1}$ based on the conditional model. This updated variable was then used in the estimations of the other variables, following the conditional model specification: $x_j^{t+1} \sim g_j(x_j \mid x_{j+1}^{t+1},...,x_{j-1}^{t+1},Z, \phi_j)$, for $j \in \{1..m\}$, where $\phi_j$ were parameters of the conditional model $g_j$. These steps were repeated for all variables $x_1,...,x_m$ and, after an initial burn-in phase, the procedure was stopped once convergence was reached.

The variables used for imputation included main outcome variables and their baseline measures, patient socio-demographics, clinical characteristics, and study design variables, such as site and intervention indicators, as well as dummy indicators for each follow-up assessment.

We used interval regressions to incorporate the theoretical bounds of the clinical outcome variables to increase the efficiency and accuracy of the imputation procedure, (e.g., we restricted the imputed values of PHQ-9 to lie within the 0 to 27 intervals). To impute the binary variables, e.g. whether the urine drug test was positive, logistic regressions were used. Other conditional models were specified to be multiple linear regressions.

After running analysis on the twenty data sets we imputed via step 1, we combined these separate estimates from the different data sets to arrive at our final estimates.
eAppendix 4. Analytical methods – data analysis

First, we compared distributions of baseline characteristics between participants who received the IIDEA intervention and participants in the usual care group, to assess the balance of the observed covariates. For each follow-up assessment, we compared the baseline covariates to detect any statistically significant baseline differences between those who completed assessment versus those who did not.

Second, to assess the effect of the intervention, we conducted an intent-to-treat (ITT) analysis, using a multilevel, multivariate regression model to assess changes in outcome variables over time in the treatment and control groups. To account for the nature of longitudinal data, the ITT analysis was carried out by fitting multilevel mixed-effects models to allow for valid variance calculation and statistical inference. The multilevel models we used included random effects at patient level to account for within-patient correlations, and robust clustered standard errors to account for within-clinic correlations due to patients nesting within same clinic. Letting $t^*$ denote the post-intervention month, we estimated the following model:

$$Y_{it} = \beta_0 + \beta_1 \text{Intervention}_i + \beta_2 \text{Time}_t + \beta_3 \text{Intervention}_i \times \text{Time}_t + \beta_4 (\text{Time}_t - t^*)$$

$$+ \beta_5 \text{Intervention}_i \times (\text{Time}_t - t^*) + \beta_6 X_i + \epsilon_{it}$$

where the patient-specific random intercept can be written as $\beta_{0i} = \alpha_{0i} + \omega_{0i}$. We fitted linear models for continuous outcomes and logistic models for binary outcomes. Although scores on the ASI drug and ASI alcohol measures range from 0 to 1, participant scores on these measures were rescaled to a range of 0 to 100 prior to regression analyses (i.e., multiplied by 100). This adjustment was made to ensure meaningful regression estimates.

In the ITT analysis, individuals were assigned to the study arm to which they were randomized. Thus, $\text{Intervention}_i$ is equal to 1 if patient $i$ is randomized to the intervention arm and 0 otherwise. $\text{Time}_t$ is a continuous measure of time capturing differences in months between assessment $t$ and assessment $t - 1$. Because the intervention ended by the time of the 6-month assessment, we centered time variables at 6-month follow-up in the following way: $\text{Time}_t$ equals to -4 for 2-month follow-up, -2 for 4-month follow-up, 0 for 6-month follow-up, and 6 for 12-month follow-up. To model the pattern of outcome changes over time, our primary analysis employed linear spline models to divide the time axis into two segments, and within each segment consider piecewise linear trends (e.g., having a different time trend before and after the intervention ended). Specially, we denote $t^*$ to be the post-intervention time trend, which equals to $(\text{Time}_t - t^*)$ if $\text{Time}_t > t^*$ and 0 otherwise. This choice of linear spline models allows for the ability to test whether the time trends differ before and after the intervention was finished and that treatment effect could attenuate over time once participants do not receive more intervention. Since the time variable is centered at 6-month follow-up, the beta coefficient on $\text{Intervention}_i$ can be interpreted as the treatment effect on outcome levels evaluated at the end of the intervention. That is, testing for the significance of $\beta_3$ tests the hypothesis that the treatment was more effective than usual care in reducing substance use disorder and/or mental health illness, as evaluated at 6 months after the baseline. The coefficient of intervention by time interaction $\beta_3$, tests whether the pattern of outcome over time would be no different between treatment and usual care groups. Similarly, $\beta_5$ tests whether changes in outcome responses over time was differential between treatment and control groups after the intervention had ended. The constant term $\alpha_{0i}$ in model (1) represents the average level of outcome variables in the control group measured at the end of the intervention. The term $\omega_{0i}$ denotes the patient-specific random effect and $\epsilon_{ij}$ denotes the residual error term. $X_i$ includes the baseline response of outcomes to control for severe differences at baseline and site indicators to adjust for variations by location. Because a successful randomized control trial balances both known and unknown confounders between treatment and control groups, our primary ITT analyses did not control for other covariates. We tested the robustness of the results by further adjusting for current medication use in a sensitivity analysis. The adjusted results did not differ from the main results (available upon request).

Our study sample is comprised of a broad clinical population, with a substantial number of patients who only had mild baseline symptoms in substance use and mental health. To test the hypothesis that intervention would be more effective among those with moderate to severe symptoms, we extended model (1) to include an intervention by baseline severity interaction. Baseline severity variable is dummy-coded, with 1 if participant’s baseline outcome score equals or exceeds a moderate level of symptoms and 0 otherwise. Thresholds of moderate severity level are:

© 2019 Alegría et al. JAMA Network Open.
0.1 for ASI-Alcohol and ASI-Drug in their original scales, being positive to at least 1 substance on the urine drug test, 10 for PHQ-9, 10 for GAD-7, 33 for PCL-5, 1.5 for HSCL-20, 35 for the composite mental score and 20 for the composite substance score. These severity thresholds are based on either published literature or on the 50th percentile of the baseline outcome scores (except for the ASI-Drug measure). Because the distribution of ASI-Drug in this sample is highly rightly-skewed, 75th percentile of the baseline outcome scores was chosen to be the severity threshold. An interaction between this baseline severity dummy and intervention was then added into Model (1). With this adjustment, $\beta_1$ now represents the intervention effect at 6-month follow-up for those with mild symptoms. Coefficient of the baseline severity and intervention interaction measures the incremental effect of the treatment among participants with moderate to severe baseline symptoms. Results were discussed with further details in the manuscript.

We also examined whether effectiveness of the intervention was moderated by other person-level factors, such as income and recent immigration status, by estimating an interaction term between person-level factor and the intervention dummy. Income status was represented by a dummy coded variable for those whose self-reported personal income less than 15,000 US dollars, with medium to high income level as the reference category. Status of recent immigration was dummy-coded, with 1 if the length of residence in the host country was less than 5 years and 0 otherwise. Results show that effectiveness of the intervention was no different with respect to family income status. In contrast, we did observe a stronger treatment effect among recent immigrants in reducing mental health symptoms and substance use (results available upon request).

In the context of multisite randomized control trial, we tested whether the intervention was equally effective at three sites. To do this, we added an intervention by site interaction to Model (1) and tested if the interaction term was statistically significant, with Boston as the reference group (see eTable 3). We found that there was no evidence of differential effectiveness by site except for PHQ-9, where the intervention effect appears to be smaller in Madrid than Boston. Additionally, we tested whether the intervention effect was the same among patients who received the most sessions by telephone vs. in-person. We recategorized participants into four mutually exclusive groups: 1) patients in the control arm; 2) intervention patients who received zero sessions, 3) intervention patients who received most sessions by telephone, and 4) intervention patients who received most sessions in-person. When there was a tie between number of sessions received by telephone vs. in-person, we randomly assigned the patient into either group. Finally, we replaced the dummy-coded intervention variable with this multi-group variable to refit the ITT model (see eTable 4). For the study Latino immigrants with co-occurring problems, treatment in person appears more effective than by phone.

For the potential moderators (i.e., baseline severity level, site and type of treatment received), we also tested whether trajectories of treatment effects over time were moderated by the potential moderator. For instance, we estimated the difference in treatment time effects between those with moderate to severe symptoms at baseline vs those with mild symptoms using the interaction between the baseline severity indicator and treatment*time indicators. The statistically insignificant interaction coefficients indicated that the trajectory of treatment effects over time was not moderated by baseline severity level. Thus, primary moderation analysis did not include these additional interactions terms.

In secondary analyses, we used dosage (defined as number of treatment sessions received) as the independent variable of interest, categorized as dosage equal to zero (control group), 0-3 sessions (inadequate treatment for the intervention group), 4 or more sessions (adequate treatment for the intervention group). While this analysis no longer relies on random assignment, it serves to provide confirmation of the results from the intent-to-treat analysis and provides further estimates of the magnitude of the intervention effects on the main outcomes evaluated at 6 months after the baseline. Because the dosage analysis relies on actual treatment received, we used the original non-imputed sample for the analysis. To check the robustness of the dosage analysis, we conducted a third sensitivity analysis with an alternative categorization of treatment dosage. This alternative dosage variable has three categories: zero (control group), 0-5 sessions (intervention group) and 6 sessions or more (intervention group). We considered completion of treatment if the patient received 6 sessions or more, as 6 sessions cover the core components of the intervention (i.e. cognitive restructuring exercises, mindfulness practice, relapse prevention, etc.). Separating those with six or more sessions will allow us to examine the treatment effect for those who completed the core components of the intervention (see eTable 2).
We also performed separate sensitivity analyses to ensure the robustness of the results to alternative modeling strategies, estimation methods and how the missing data was handled.

Our first set of sensitivity analyses explored alternative methods for handling baseline response. Our primary analysis was carried out through analysis of covariance, which analyzes post-baseline responses, and makes an adjustment for the baseline response by including it as a covariate. In the sensitivity analysis, we retained the baseline response as part of the outcome vector and assumed the group means were equal at baseline, as is appropriate in a randomized control trial. The analytical data was extended to a longitudinal dataset, consisting of 5 repeated measurements per person, with baseline assessment included as an additional repeated assessment. To proceed, we re-estimated model (1) with this new dataset. Since baseline was added as additional time point, the $Time_i$ variable in this analysis, equals to -6 for baseline, -4, -2, 0 and 6 for the two, four, six and twelve-month assessments. Results of this sensitivity analysis are available upon request. We chose to present analysis of covariance as our primary results due to its potential efficiency gain. Because the baseline value has been obtained before any study intervention, i.e., the mean response that baseline is independent of treatment assignment, adjustment for baseline through covariance analysis will be more efficient, as it yields estimates of treatment effects with smaller standard errors.

Our primary ITT analysis modeled the nonlinear trend by fitting piecewise linear trends before and after treatment completion. We explored alternative modeling strategy in sensitivity analysis where the post-treatment trend in model (1) was replaced with a quadratic term of time trend to model the non-linear time trend and found similar results. Linear spline models were finally chosen because they generally provide a flexible way to accommodate many non-linear trends that cannot be approximated by simple polynomials in time.

Next, we checked the robustness of the results with respect to estimation method. Our primary ITT analysis used multilevel mixed-effect models to account for inter-participant correlation due to repeated measurement. Our sensitivity analysis instead used generalized estimating equations (GEE). GEE estimation will produce unbiased estimates if the underlying correlation structure is correctly specified. However, it has the limitation of being less efficient than a properly specified mixed model and does not allow weights to vary within clusters/panels.

We also performed a sensitivity analysis using a propensity score weighting approach to balance the treatment and control groups on residual differences within each site after randomization. To do so, we estimated a logistic regression model of membership in the treatment group conditional on baseline measures and interactions between baseline measures and site and generated the predicted probability ($phat$) of being assigned to the treatment group. In multivariate regression models, treatment group participants were given a weight of $1/phat$ whereas control group participants were given a weight of $1/1-phat$ to balance the groups on their predicted probability of being in the treatment group. Propensity score weights were then applied to all regression analysis. The weighted results remain similar to our main results.

Finally, we re-ran the dosage analysis with imputed data as a sensitivity check. In comparison to list-wise deletion, estimates of the imputed data from multiple imputation have slightly larger effect sizes (available from the authors upon request). This is because the intervention effects were greater among patients with moderate to severe symptomatology compared to those with mild symptomatology. Patients with missing observations, i.e., those who did not complete all follow-up assessments, had higher baseline symptom severity (eTable 5). Thus, by including these patients when using multiple imputation, the estimated intervention effect was amplified. Our primary analysis, which used non-imputed data and list-wise deletion, provides a conservative estimate of the actual effect of the treatment dosage.
eFigure 1. IIDEA study flow

**Approach**
Conducted by research assistants [3785 Latino migrants approached across 3 sites] recruited in primary care and community agencies in Boston [1476], Madrid [1282], and Barcelona.

- Declined to participate [806]
- Did not respond to “call back” [695]

**Eligibility screening**
Research assistants conducted [2284 potential participants] consent, admission, and short screener.

**Inclusion criteria:** 18-70 years old, self-identified Latinos, positive screen to mental health and substance use problems by using the AC-OK co-occurring disorder screener.

**Exclusion criteria:** Current mental health or substance treatment in the last 3 months or any scheduled visit to receive treatment in the next month, inability to demonstrate capacity to consent to the study and evidence of suicidal risk or harm in the last 30 days.

**Enrollment**
Conducted by research assistants [384 eligible participants] baseline and urine drug test.

**Randomization**
[341 randomly assigned and included in intent to treat analysis]

- Lost to follow-up [37]
- Declined to participate [6]

**Intervention group** [172]
Received 10-12 IIDEA therapy sessions by trained clinicians.

- Follow up assessment [130] and urine drug test 2 months after baseline
- Follow up assessment [119] and urine drug test 4 months after baseline
- Follow up assessment [128] and urine drug test 6 months after baseline
- Follow up assessment [131] and urine drug test 12 months after baseline

**Control group** [169]
Receive four usual

- Follow up assessment [134] and urine drug test 2 months after baseline
- Follow up assessment [123] and urine drug test 4 months after baseline
- Follow up assessment [139] and urine drug test 6 months after baseline
- Follow up assessment [132] and urine drug test 12 months after baseline

HIV testing [246]
STI testing [245]
eFigure 2. Heterogenous treatment effects on primary and secondary outcomes across different subgroups

Panel A. Participants with Mild vs. Moderate to Severe Symptoms at Baseline

Note: Effect sizes are calculated in the metric of Odds Ratios for binary outcomes and in the metric of Cohen’s d for continuous outcomes. Positive drug test is a binary outcome based on the results of urine drug tests for use of any six drug types (amphetamines, benzodiazepines, cocaine, methamphetamine, heroin and marijuana). ASI: Addiction Severity Index; PHQ: Patient Health Questionnaire; GAD: Generalized Anxiety Disorder; PCL: Post-Traumatic Stress Disorder Checklist; HSCL: Hopkins Symptom Checklist; DAST: Drug Abuse Screening Test; AUDIT: Alcohol Use Disorders Identification Test; BDEPQ: Benzodiazepine Dependence Questionnaire.
# eTable 1. List of Assessment Measures/Variables

<table>
<thead>
<tr>
<th>Socio-Demographic variables</th>
<th>Site (Boston/Madrid/Barcelona), age (18-34, 35-49, 50+), gender (male/female), race/ethnicity (White, Black, Indigenous/Native American, Hispanic/Latino/Caribbean, Mixed), education level (less than high school/HS Diploma, GED, vocational school, or more), Total Personal Income Last Year (&lt; 15000 US Dollars or ≥ 15000 US Dollars)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clinical Measures</strong></td>
<td></td>
</tr>
<tr>
<td>Depression - Patient Health Questionnaire (PHQ-9)</td>
<td>9-item screening questionnaire to determine severity of depressive symptoms. Internal consistency was $\alpha = 0.85$. Spanish version has good agreement with independent mental health professional Dx ($k = 0.74$; overall accuracy, $88%$; sensitivity, $87%$; specificity, $88%$) (27).</td>
</tr>
<tr>
<td>Generalized Anxiety - General Anxiety Disorder 7-item screener (GAD-7)</td>
<td>7-item screening tool and severity measure for generalized anxiety disorder. In Spanish, a cut-off point of 10 showed adequate values for sensitivity (86.8%) and specificity (93.4%); AUC statistically significant [AUC = 0.957-0.985; $p &lt; 0.001$]; Internal consistency was $\alpha = 0.86$ (28).</td>
</tr>
<tr>
<td>Post-traumatic Stress Disorder - PTSD Checklist for DSM-5 (PCL-5)</td>
<td>The PCL-5 is a 20-item questionnaire, corresponding to the DSM-5 symptom criteria for PTSD. We consider 33 as a cutoff given the latest data. The internal consistency was $\alpha = 0.94$ (29).</td>
</tr>
<tr>
<td>Drug Abuse - Drug Abuse Screening Test (DAST-10)</td>
<td>10-item self-report instrument that has been condensed from the 28-item DAST. It is designed for clinical screening of substance use. The DAST-10 yields a quantitative index of the degree of consequences related to drug abuse. ($\alpha = 0.87$) (30).</td>
</tr>
<tr>
<td>Alcohol Abuse – Alcohol Use Disorders Identification Test – C (AUDIT-C)</td>
<td>World Health Organization screener for excessive drinking. 3-question screen that evaluates lifetime and past 30-day alcohol use behaviors ($\alpha = 0.78$) (31).</td>
</tr>
<tr>
<td>Benzodiazepines - Benzodiazepine Dependence Questionnaire (BDEPQ)</td>
<td>The BDEPQ is designed to measure dependence on benzodiazepine tranquilizers, sedatives, and hypnotics. We selected 10 representative items from this questionnaire. ($\alpha = 0.89$) (32).</td>
</tr>
<tr>
<td>Alcohol – Addiction Severity Index Lite (ASI Lite) – Alcohol and Drug - Addiction Severity Index Lite (ASI Lite) - Drug</td>
<td>Shortened version of the Addiction Severity Index (ASI) which obtains lifetime information about problem behaviors, as well as problems within the previous 30 days. We used drug and alcohol questions only ($\alpha=0.84$ for alcohol and $\alpha=0.70$ for drugs) (33).</td>
</tr>
<tr>
<td>HSCL-20 – Hopkins Symptom Checklist-20</td>
<td>The HSCL-20 is a symptom inventory for depression that indicates overall mental health symptoms. Internal consistency was $\alpha=0.94$ (34).</td>
</tr>
<tr>
<td>Mindfulness Awareness - Mindfulness Awareness Attention Scale</td>
<td>The MAAS is a 15-item scale designed to assess a core characteristic of dispositional mindfulness, namely, open or receptive awareness of and attention to what is taking place in the present. The scale shows strong</td>
</tr>
</tbody>
</table>
psychometric properties and has been validated with community samples. (α = 0.87) (35, 36).

| Smoking - Fagerstrom Test for Nicotine Dependence | The Fagerstrom Test is a standard instrument for assessing the intensity of physical addiction to nicotine, and designed to provide an ordinal measure of nicotine dependence related to cigarette smoking. It contains six items that evaluate the quantity of cigarette consumption, the compulsion to use, and dependence (α =0.75) (37). |
| Trauma Exposure - Brief Trauma Questionnaire (BTQ) | 10-item self-report measure that examines experiences with potentially traumatic events that would meet Criterion A (serious injury/life threat/subjective response) for PTSD diagnosis according to DSM-IV. It is derived from the Brief Trauma Interview. Internal consistency was (38). |
| Any positive in Urine Test | Using the DrugCheck for drug metabolites with a binary outcome (yes/no) for use of any of this six drug types: amphetamines, benzodiazepines, cocaine, methamphetamine, marijuana, and heroin |

### Cultural/Social Factors

| Citizenship (Yes/No) | Immigration status was defined by whether the participant was a citizen or non-citizen. |
| Years in US/Spain | |
| Number of Home Visits in the Past 12 Months | |
| Discrimination Scale | Everyday Discrimination Scale (39) (α=0.82). The scale measures chronic and routine unfair treatment in everyday life. Adopted from the Detroit Area Study and National Latino and Asian-American Study (NLAAS). Participants were asked 9 questions pertaining to discrimination that affect Latinos and other minorities. |
| Ethnic Identity Scale | The 3-item Ethnic Identity Scale (α = 0.73), derived from the 35-item Cultural Identity Scale for Latino Adolescents (40), was used to gauge cultural identity. Questions ask about what culture or ethnic/racial groups an individual identifies with. |
| Intercultural/Family Conflict Scale | We used 3 items of the Family/Culture Stress subscale of the Hispanic Stress Inventory (HIS) (α=0.66), designed to measure family interference with personal goals, arguments with family members, and the breakdown of the family unit (41-43). |
| Sense of Belonging | One question derived from qualitative research from the research team in previous studies (44). The question is: You feel like you don’t belong either in your Latino country or in the U.S.? |
Recruitment Site

Recruitment Site
Primary Care Clinics/Community Sites/Emergency Room/Patient Referrals

CHANGES FROM THE ORIGINAL PROTOCOL:

We updated measures from the originally registered protocol: We replaced the Posttraumatic Cognitions Inventory with the PTSD Checklist for DSM-5 to follow the most updated DSM-V guidelines. We also added the Hopkins Symptoms Checklist after the registration of the trial to provide information about overall mental health symptoms/psychological distress. We replaced our originally registered follow-up assessment schedule (1, 3, 6 months) with 2, 4, 6, and 12-month assessments, to better align with intervention flow, at the recommendation of our Advisory Board.
### eTable 2. Analysis of participation rates (receiving 6+ sessions) in IIDEA intervention on main outcomes at 6-month follow-up (N=341)\(^a\)

<table>
<thead>
<tr>
<th></th>
<th>ASI-Drug</th>
<th>ASI-Alcohol</th>
<th>Positive in Urine Drug Test</th>
<th>PHQ-9</th>
<th>GAD-7</th>
<th>PCL-5</th>
<th>HSCL-20</th>
<th>Composite Mental Score</th>
<th>Composite Substance Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention Patients with 0-5 Sessions</td>
<td>1.09</td>
<td>0.97</td>
<td>0.01</td>
<td>-0.92</td>
<td>0.33</td>
<td>-1.90</td>
<td>-0.07</td>
<td>-1.16</td>
<td>2.43</td>
</tr>
<tr>
<td></td>
<td>(0.87)</td>
<td>(1.44)</td>
<td>(0.47)</td>
<td>(0.55)</td>
<td>(0.67)</td>
<td>(1.45)</td>
<td>(0.06)</td>
<td>(2.05)</td>
<td>(1.73)</td>
</tr>
<tr>
<td>Intervention Patients with 6+ Sessions</td>
<td>-1.50*</td>
<td>0.61</td>
<td>-0.83</td>
<td>-1.70*</td>
<td>-1.36*</td>
<td>-4.74*</td>
<td>-0.30***</td>
<td>-6.40*</td>
<td>-4.30**</td>
</tr>
<tr>
<td></td>
<td>(0.67)</td>
<td>(1.60)</td>
<td>(0.62)</td>
<td>(0.74)</td>
<td>(0.58)</td>
<td>(2.14)</td>
<td>(0.11)</td>
<td>(2.42)</td>
<td>(1.25)</td>
</tr>
<tr>
<td>Madrid(^b)</td>
<td>-1.11</td>
<td>-2.22</td>
<td>-0.35</td>
<td>-0.18</td>
<td>-0.36</td>
<td>-2.15</td>
<td>-0.01</td>
<td>-1.64</td>
<td>1.78</td>
</tr>
<tr>
<td></td>
<td>(1.15)</td>
<td>(2.56)</td>
<td>(0.59)</td>
<td>(0.73)</td>
<td>(0.72)</td>
<td>(1.73)</td>
<td>(0.14)</td>
<td>(2.39)</td>
<td>(1.75)</td>
</tr>
<tr>
<td>Barcelona</td>
<td>-1.18</td>
<td>-0.61</td>
<td>-0.52</td>
<td>-1.46*</td>
<td>-1.39**</td>
<td>-5.48***</td>
<td>-0.18</td>
<td>-5.76**</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>(0.83)</td>
<td>(1.55)</td>
<td>(0.36)</td>
<td>(0.66)</td>
<td>(0.41)</td>
<td>(1.22)</td>
<td>(0.10)</td>
<td>(1.68)</td>
<td>(1.09)</td>
</tr>
<tr>
<td>Baseline Outcome</td>
<td>0.62***</td>
<td>0.42***</td>
<td>2.25***</td>
<td>0.47***</td>
<td>0.45***</td>
<td>0.51***</td>
<td>0.48***</td>
<td>0.53***</td>
<td>0.55***</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.05)</td>
<td>(0.36)</td>
<td>(0.05)</td>
<td>(0.06)</td>
<td>(0.05)</td>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Constant</td>
<td>1.61</td>
<td>4.79*</td>
<td>-1.49***</td>
<td>3.64***</td>
<td>2.69***</td>
<td>9.25***</td>
<td>0.42**</td>
<td>10.20**</td>
<td>3.64*</td>
</tr>
<tr>
<td></td>
<td>(0.94)</td>
<td>(2.07)</td>
<td>(0.32)</td>
<td>(0.84)</td>
<td>(0.73)</td>
<td>(1.69)</td>
<td>(0.14)</td>
<td>(2.92)</td>
<td>(1.77)</td>
</tr>
<tr>
<td>N</td>
<td>268</td>
<td>268</td>
<td>257</td>
<td>268</td>
<td>268</td>
<td>268</td>
<td>268</td>
<td>268</td>
<td>268</td>
</tr>
</tbody>
</table>

\(^a\) This analysis only uses data from the 6-month follow up assessment. Missing data is list-wise deleted. The analytical sample for each regression slightly varies as the missing pattern of each outcome variable slightly differs.

\(^b\) The reference group is patients in control arm and the reference site is Boston.

Standard errors are reported in parentheses. Coefficients in terms of the log odds are reported for binary outcome.

*= p<0.05, **=p<0.01, ***=p<0.001.

ASI: Addiction Severity Index; PHQ: Patient Health Questionnaire; GAD: Generalized Anxiety Disorder; PCL: Post-Traumatic Stress Disorder Checklist; HSCL: Hopkins Symptom Checklist; DAST: Drug Abuse Screening Test; AUDIT: Alcohol Use Disorders Identification Test; BDEPQ: Benzodiazepine Dependence Questionnaire.
### eTable 3. Intent-to-treat analysis of IIDEA intervention on main outcomes by site (N=341) 

<table>
<thead>
<tr>
<th></th>
<th>ASI-Drug</th>
<th>ASI-Alcohol</th>
<th>Positive in Urine Drug Test</th>
<th>PHQ-9</th>
<th>GAD-7</th>
<th>PCL-5</th>
<th>HSCL-20</th>
<th>Composite Mental Health</th>
<th>Composite Substance Use</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intervention</strong></td>
<td>-0.83</td>
<td>0.15</td>
<td>-0.70</td>
<td>-2.04*</td>
<td>-1.85</td>
<td>-6.00*</td>
<td>-0.32</td>
<td>-7.93*</td>
<td>-2.37</td>
</tr>
<tr>
<td></td>
<td>(1.71)</td>
<td>(0.63)</td>
<td>(0.90)</td>
<td>(1.00)</td>
<td>(2.46)</td>
<td>(0.14)</td>
<td>(3.25)</td>
<td>(2.62)</td>
<td></td>
</tr>
<tr>
<td><strong>Time</strong></td>
<td>-0.13</td>
<td>-0.97**</td>
<td>0.01</td>
<td>-0.38**</td>
<td>-0.43**</td>
<td>-1.36**</td>
<td>-0.06**</td>
<td>-1.72**</td>
<td>-0.52**</td>
</tr>
<tr>
<td></td>
<td>(0.16)</td>
<td>(0.10)</td>
<td>(0.11)</td>
<td>(0.10)</td>
<td>(0.32)</td>
<td>(0.02)</td>
<td>(0.37)</td>
<td>(0.26)</td>
<td></td>
</tr>
<tr>
<td><strong>Intervention*Time</strong></td>
<td>0.11</td>
<td>0.46</td>
<td>-0.05</td>
<td>-0.05</td>
<td>0.10</td>
<td>0.44</td>
<td>-0.01</td>
<td>0.28</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.39)</td>
<td>(0.15)</td>
<td>(0.13)</td>
<td>(0.50)</td>
<td>(0.02)</td>
<td>(0.53)</td>
<td>(0.40)</td>
<td></td>
</tr>
<tr>
<td><em><em>(Time-t</em>)</em>*</td>
<td>0.10</td>
<td>0.96*</td>
<td>-0.04</td>
<td>0.25</td>
<td>0.42**</td>
<td>1.02*</td>
<td>0.06**</td>
<td>1.40**</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td>(0.21)</td>
<td>(0.42)</td>
<td>(0.14)</td>
<td>(0.14)</td>
<td>(0.41)</td>
<td>(0.02)</td>
<td>(0.51)</td>
<td>(0.34)</td>
<td></td>
</tr>
<tr>
<td><em><em>Intervention</em>(Time-t</em>)**</td>
<td>-0.12</td>
<td>-0.53</td>
<td>0.10</td>
<td>0.21</td>
<td>-0.08</td>
<td>-0.11</td>
<td>0.02</td>
<td>0.09</td>
<td>-0.17</td>
</tr>
<tr>
<td></td>
<td>(0.26)</td>
<td>(0.52)</td>
<td>(0.22)</td>
<td>(0.19)</td>
<td>(0.64)</td>
<td>(0.03)</td>
<td>(0.71)</td>
<td>(0.53)</td>
<td></td>
</tr>
<tr>
<td><strong>Madrid</strong></td>
<td>-0.90</td>
<td>-0.49</td>
<td>-1.22*</td>
<td>-0.95</td>
<td>-1.21</td>
<td>-5.21**</td>
<td>-0.17</td>
<td>-5.09*</td>
<td>1.38</td>
</tr>
<tr>
<td></td>
<td>(1.33)</td>
<td>(1.87)</td>
<td>(0.51)</td>
<td>(0.74)</td>
<td>(1.75)</td>
<td>(0.11)</td>
<td>(2.27)</td>
<td>(2.09)</td>
<td></td>
</tr>
<tr>
<td><strong>Barcelona</strong></td>
<td>-0.76</td>
<td>-0.20</td>
<td>-0.34</td>
<td>-0.94</td>
<td>-1.62**</td>
<td>-5.45**</td>
<td>-0.20</td>
<td>-5.35**</td>
<td>1.44</td>
</tr>
<tr>
<td></td>
<td>(1.22)</td>
<td>(1.23)</td>
<td>(0.52)</td>
<td>(0.61)</td>
<td>(1.71)</td>
<td>(0.10)</td>
<td>(1.97)</td>
<td>(1.77)</td>
<td></td>
</tr>
<tr>
<td><strong>Intervention*Madrid</strong></td>
<td>1.14</td>
<td>-0.08</td>
<td>1.32</td>
<td>1.89*</td>
<td>1.85</td>
<td>5.07</td>
<td>0.24</td>
<td>7.34*</td>
<td>2.15</td>
</tr>
<tr>
<td></td>
<td>(1.72)</td>
<td>(2.49)</td>
<td>(0.78)</td>
<td>(0.95)</td>
<td>(2.95)</td>
<td>(0.14)</td>
<td>(3.53)</td>
<td>(2.79)</td>
<td></td>
</tr>
<tr>
<td><strong>Intervention*Barcelona</strong></td>
<td>1.05</td>
<td>-0.28</td>
<td>0.08</td>
<td>0.76</td>
<td>1.31</td>
<td>2.52</td>
<td>0.12</td>
<td>4.10</td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td>(1.63)</td>
<td>(1.69)</td>
<td>(0.67)</td>
<td>(0.87)</td>
<td>(2.41)</td>
<td>(0.13)</td>
<td>(3.02)</td>
<td>(2.68)</td>
<td></td>
</tr>
<tr>
<td><strong>Baseline Outcome</strong></td>
<td>0.56***</td>
<td>0.47***</td>
<td>3.55***</td>
<td>0.53***</td>
<td>0.49***</td>
<td>0.55**</td>
<td>0.45***</td>
<td>0.58***</td>
<td>0.51***</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.05)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td></td>
</tr>
<tr>
<td><strong>Constant</strong></td>
<td>2.39</td>
<td>6.83***</td>
<td>-2.24***</td>
<td>3.19***</td>
<td>3.31***</td>
<td>10.49**</td>
<td>0.58***</td>
<td>10.89***</td>
<td>6.18**</td>
</tr>
<tr>
<td></td>
<td>(1.27)</td>
<td>(1.48)</td>
<td>(0.50)</td>
<td>(0.67)</td>
<td>(1.95)</td>
<td>(0.11)</td>
<td>(2.62)</td>
<td>(2.00)</td>
<td></td>
</tr>
</tbody>
</table>

---

*This analysis uses longitudinal data of 341 participants with four follow-up assessments per participant. Each outcome variable was measured four times at 2 months, 4 months, 6 months and 12 months follow-up, respectively. The unit observation is a specific follow-up assessment. Missing data is handled by multiple imputation method (eAppendix 3, supplement).

**Time** is a continuous variable which equals to -4, -2, 0 and 6 for RA5 to RA5, respectively.

---

© 2019 Alegria et al. JAMA Network Open.
### Table 4. Intent-to-treat analysis of IIDEA intervention on main outcomes by received intervention type (N=341)

<table>
<thead>
<tr>
<th></th>
<th>ASI-Drug</th>
<th>ASI-Alcohol</th>
<th>Positive in Urine Drug Test</th>
<th>PHQ-9</th>
<th>GAD-7</th>
<th>PCL-5</th>
<th>HSCL-20 Composite Mental Score</th>
<th>Composite Substance Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention Patients with 0 sessions (ref=Control Patients)(b)</td>
<td>1.80**</td>
<td>2.31*</td>
<td>0.46</td>
<td>0.11</td>
<td>0.71</td>
<td>0.86</td>
<td>0.03</td>
<td>1.88</td>
</tr>
<tr>
<td></td>
<td>(0.56)</td>
<td>(1.13)</td>
<td>(0.50)</td>
<td>(0.44)</td>
<td>(0.46)</td>
<td>(1.20)</td>
<td>(0.06)</td>
<td>(1.49)</td>
</tr>
<tr>
<td>Intervention Patients with Most Sessions Received by Telephone</td>
<td>0.08</td>
<td>-1.65</td>
<td>0.02</td>
<td>-0.33</td>
<td>-0.43</td>
<td>-4.71*</td>
<td>-0.10</td>
<td>-3.35+</td>
</tr>
<tr>
<td></td>
<td>(0.71)</td>
<td>(1.84)</td>
<td>(0.53)</td>
<td>(0.57)</td>
<td>(0.52)</td>
<td>(1.87)</td>
<td>(0.10)</td>
<td>(1.96)</td>
</tr>
<tr>
<td>Intervention Patients with Most Sessions Received in Person</td>
<td>-1.07*</td>
<td>-1.85*</td>
<td>-0.63*</td>
<td>-1.45***</td>
<td>-1.61***</td>
<td>-5.06***</td>
<td>-0.29***</td>
<td>-6.41***</td>
</tr>
<tr>
<td>Time</td>
<td>(0.50)</td>
<td>(0.94)</td>
<td>(0.28)</td>
<td>(0.42)</td>
<td>(0.44)</td>
<td>(1.33)</td>
<td>(0.07)</td>
<td>(1.59)</td>
</tr>
<tr>
<td></td>
<td>-0.05</td>
<td>-0.75**</td>
<td>-0.01</td>
<td>-0.42***</td>
<td>-0.39***</td>
<td>-1.17***</td>
<td>-0.07***</td>
<td>-1.62***</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.27)</td>
<td>(0.06)</td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.20)</td>
<td>(0.01)</td>
<td>(0.24)</td>
</tr>
<tr>
<td></td>
<td>0.01</td>
<td>0.72*</td>
<td>-0.00</td>
<td>0.39***</td>
<td>0.39***</td>
<td>1.02***</td>
<td>0.07***</td>
<td>1.53***</td>
</tr>
<tr>
<td></td>
<td>(0.13)</td>
<td>(0.33)</td>
<td>(0.09)</td>
<td>(0.11)</td>
<td>(0.09)</td>
<td>(0.28)</td>
<td>(0.01)</td>
<td>(0.34)</td>
</tr>
<tr>
<td>Madrid(b)</td>
<td>-1.11</td>
<td>-2.22</td>
<td>-0.35</td>
<td>-0.18</td>
<td>-0.36</td>
<td>-2.15</td>
<td>-0.01</td>
<td>-1.64</td>
</tr>
<tr>
<td></td>
<td>(1.15)</td>
<td>(2.56)</td>
<td>(0.59)</td>
<td>(0.73)</td>
<td>(0.72)</td>
<td>(1.73)</td>
<td>(0.14)</td>
<td>(2.39)</td>
</tr>
<tr>
<td>Barcelona</td>
<td>-1.18</td>
<td>-0.61</td>
<td>-0.52</td>
<td>-1.46*</td>
<td>-1.39**</td>
<td>-5.48***</td>
<td>-0.18</td>
<td>-5.76**</td>
</tr>
<tr>
<td></td>
<td>(0.83)</td>
<td>(1.55)</td>
<td>(0.36)</td>
<td>(0.66)</td>
<td>(0.41)</td>
<td>(1.22)</td>
<td>(0.10)</td>
<td>(1.68)</td>
</tr>
<tr>
<td>Baseline Outcome</td>
<td>0.62***</td>
<td>0.42***</td>
<td>2.25***</td>
<td>0.47***</td>
<td>0.45***</td>
<td>0.51***</td>
<td>0.48***</td>
<td>0.53***</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.05)</td>
<td>(0.36)</td>
<td>(0.05)</td>
<td>(0.06)</td>
<td>(0.05)</td>
<td>(0.06)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Constant</td>
<td>1.61</td>
<td>4.79*</td>
<td>-1.49***</td>
<td>3.64***</td>
<td>2.69***</td>
<td>9.25***</td>
<td>0.42**</td>
<td>10.20**</td>
</tr>
<tr>
<td></td>
<td>(0.94)</td>
<td>(2.07)</td>
<td>(0.32)</td>
<td>(0.84)</td>
<td>(0.73)</td>
<td>(1.69)</td>
<td>(0.14)</td>
<td>(2.92)</td>
</tr>
</tbody>
</table>

\(a\) This analysis uses longitudinal data of 341 participants with four follow-up assessments per participant. Each outcome variable was measured four times at 2 months, 4 months, 6 months and 12 months follow-up, respectively. The unit observation is a specific follow-up assessment. Missing data is handled by multiple imputation method (eAppendix 3, supplement).

\(b\) The reference group is patients in control arm and the reference site is Boston.

Standard errors are reported in parentheses. Coefficients in terms of the log odds are reported for binary outcome.

\(*=p<0.05; **=p<0.01; ***=p<0.001.\)

ASI: Addition Severity Index; PHQ: Patient Health Questionnaire; GAD: Generalized Anxiety Disorder; PCL: Post-Traumatic Stress Disorder Checklist; HSCL: Hopkins Symptom Checklist; DAST: Drug Abuse Screening Test; AUDIT: Alcohol Use Disorders Identification Test; BDEPQ: Benzodiazepine Dependence Questionnaire.
**eTable 5. Statistically Significant baseline differences* between participants with missing/incomplete assessments and participants with complete assessments (N=341)**

<table>
<thead>
<tr>
<th></th>
<th>Participants with Complete Assessments (N=197)</th>
<th>Participants with Incomplete Assessments (N=144)</th>
<th>( \chi^2 ) or t Test</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAST, Mean (SD)</td>
<td>1.15 (2.11)</td>
<td>1.65 (2.52)</td>
<td>-1.99</td>
<td>0.048</td>
</tr>
<tr>
<td>Education Level, No. (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than High School</td>
<td>60 (30.5)</td>
<td>71 (49.3)</td>
<td>12.49</td>
<td>0.001</td>
</tr>
<tr>
<td>High School Diploma,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General Education Diploma,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocational School, or More</td>
<td>137 (69.5)</td>
<td>73 (50.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recruitment Site, No. (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary Care Clinics</td>
<td>77 (39.1)</td>
<td>71 (49.3)</td>
<td>8.55</td>
<td>0.036</td>
</tr>
<tr>
<td>Community Sites</td>
<td>56 (28.4)</td>
<td>22 (15.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emergency Room</td>
<td>11 (5.6)</td>
<td>8 (5.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient Referrals</td>
<td>53 (26.9)</td>
<td>43 (29.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Children, No. (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>138 (70.1)</td>
<td>85 (59.0)</td>
<td>4.47</td>
<td>0.035</td>
</tr>
<tr>
<td>One and More</td>
<td>59 (29.9)</td>
<td>59 (41.0)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*All baseline characteristics reported in the manuscript Table 1 are compared across the two groups. Only ones with statistically significant differences are reported.*

© 2019 Alegría et al. *JAMA Network Open.*
References


23. Stata Statistical Software: Release 14 [computer program]. College Station, TX: StataCorp LP; 2015.