Randomized Cost-Effectiveness Trial of Group Interpersonal Psychotherapy (IPT) for Prisoners with Major Depression

Jennifer E. Johnson[*],

Division of Public Health, Michigan State University College of Human Medicine, 200 East 1st St Room 366, Flint, MI 48502. Jennifer.Johnson@hc.msu.edu

Robert L. Stout,

Decision Sciences Institute, 1005 Main Street Unit 8120, Pawtucket, RI 02860. Stout@pire.org

Ted R. Miller,

Pacific Institute for Research and Evaluation, 11720 Beltsville Drive, Suite 900, Calverton, MD 20705; School of Public Health, Curtin University, Perth, Australia. miller@pire.org

Caron Zlotnick,

Butler Hospital and Brown University, 345 Blackstone Blvd, Providence, RI 02906; University of Cape Town, South Africa. Caron_Zlotnick@brown.edu

Louis A. Cerbo,

Rhode Island Department of Corrections, 39 Howard Avenue, Cranston, RI 02920. Louis.Cerbo@doc.ri.gov

Joel T. Andrade,

MHM Services, Inc. 110 Turnpike Road, Suite 308, Westborough, MA 01581. jandrade@mhm-services.com

Jessica Nargiso,

Massachusetts General Hospital, 6th floor, 151 Merrimac St, Boston MA 02114. jnargiso@partners.org

Joseph Bonner, and

Division of Public Health, Michigan State University College of Human Medicine, 200 East 1st St, Flint, MI 48502. BonnerJ5@msu.edu

Shannon Wiltsey-Stirman

National Center for PTSD, Dissemination and Training Division, 795 Willow Road (NC-PTSD 334), Menlo Park, CA 94025. Shannon.Wiltsey-Stirman@va.gov

Abstract

Objective.—This study tested the effectiveness and cost-effectiveness of interpersonal psychotherapy (IPT) for major depressive disorder (MDD) among prisoners. It is the first fully-powered randomized trial of any treatment (pharmacological or psychosocial) targeting MDD among incarcerated individuals.

[*]Corresponding author.

Method.—181 male (*n*=117) and female (*n*=64) prisoners from prison facilities in 2 states were randomized to group IPT (delivered by masters-level and nonspecialist prison counselors) for MDD plus prison treatment as usual (TAU) or to TAU alone. Participants' average age was 39 (range 20–61); 20% were African-American and 19% were Hispanic. Outcomes assessed at post-treatment and 3-month follow-up included depressive symptoms (primary; assessed using the Hamilton Rating Scale for Depression), suicidality (assessed with the Beck Scale for Suicide Ideation and Beck Hopelessness Scale), in-prison functioning (i.e., enrollment in correctional programs; discipline reports; aggression/victimization; and social support), remission from MDD, and posttraumatic stress disorder (PTSD) symptoms.

Results.—IPT reduced depressive symptoms, hopelessness, and PTSD symptoms, and increased rates of MDD remission relative to prison TAU alone. Effects on hopelessness were particularly strong. Cost per patient was $2054 including costs for IPT training and supervision, and $575 without these costs. For providers running their second or subsequent IPT group, cost per additional week in remission from MDD (relative to TAU alone) was $524 ($148 excluding training and supervision costs, which would not be needed for established programs).

Conclusions.—IPT is effective and cost-effective and we recommend its use for MDD among prisoners. It is currently the only treatment for MDD evaluated among incarcerated individuals.

More than 2.3 million people are held in United States prisons and jails on any given day (Carson, 2015; Minton, 2015). Incarcerated individuals have high rates of mental health disorders (James and Glaze, 2006). Major depressive disorder (MDD) is the most common serious mental illness among incarcerated individuals (Fazel and Danesh, 2002). National surveys have found that up to 23.5% of prisoners met criteria for MDD within the past 12 months (James and Glaze, 2006). MDD carries a high death and disability burden (Ferrari et al., 2013), with serious consequences for prisoners. In-prison effects of MDD include risk for suicide (Baillargeon et al., 2009), dropout from correctional treatment programs (Brady, Krebs, and Laird, 2004; Hiller, Knight, & Simpson, 1999), physical victimization by other inmates (Blitz, Wolff, and Shi, 2008), inability to assertively protect oneself (Varese, Pelowski, Riedel, and Heiby, 1998), rejection by other inmates (Marcus, Hamlin, and Lyons, 2001), and aggressive acting out (Varese et al., 1998). MDD is also associated with substance use relapse and with return to correctional custody once individuals are released from prison (Baillargeon et al., 2009). The impairment in occupational, social, and family functioning seen with MDD can add to the challenges individuals leaving prison face as they try to re-integrate into their communities.

Despite the prevalence and consequences of MDD, there has never been a fully powered randomized controlled trial (RCT) of any treatment (pharmacological or psychosocial) for MDD in an incarcerated population. In contrast to the thousands of RCTs of treatments for individuals with MDD in the community (150 published in 2007 alone; Weinberger, McKee, and Mazure, 2010), the largest previous RCT of MDD for prisoners had a sample size of 38 (Johnson and Zlotnick, 2012). The current study evaluating group interpersonal

psychotherapy (IPT) for MDD among men and women prisoners in two states is the first fully-powered RCT of any treatment for MDD in any incarcerated population.

IPT is an evidence-based treatment for MDD with proven effectiveness in non-criminal justice settings (Elkin, 1989; National Collaborating Centre for Mental Heath, 2010). IPT identifies a current interpersonal crisis in one of four areas (an interpersonal dispute, a change in life circumstances, grief, or social isolation) as the proximal trigger for the current depressive episode and addresses it by helping the individual to improve communication, change relationship expectations, or adapt to changes within the context of building or better utilizing a social support network (Weissman, Markowitz, and Klerman, 2000). This approach is relevant for prisoners because they experience unusually high numbers of stressful and disruptive life events and conflicts with family, friends, and incarcerated peers (Keaveny and Zauszniewski, 1999; Klein and Santiago, 2003). In fact, one study found that 90% of newly sentenced prisoners experienced clinically significant levels of life change and loss stressors in the past year, and that their *average* "life change unit" score was more than twice the clinical cutoff (Keaveny and Zauszniewksi, 1999). Addressing grief is also relevant for prisoners. Keaveny and Zauszniewski found that 26% of newly sentenced prisoners had lost a family member to death in the past year, and 27% had experienced the death of a close friend during that time. Bereavement may be exacerbated by other losses such as family dissolution and loss of parental rights to children. Prisoners also suffer from social isolation, and often have difficulty trusting others and problems establishing boundaries, created by interpersonal histories that have often included physical and/or sexual abuse (Debowska and Boduszek, 2017; McDaniels-Wilson and Belknap, 2008). Social support can buffer stressors and aid in recovery from MDD. In addition, strategies, such as IPT, that build social support can aid in successful completion of prison time, successful completion of correctional programming, and successful community re-entry (Parsons and Warner-Robbins, 2002; Skeem, Louden, Manchak, Haddad, and Vidal, 2009). IPT's acceptability across a wide variety of cultures and effectiveness in a group format (Bolton et al., 2003) give it good potential for uptake within the prison system.

Because the criminal justice system has a primary public safety mandate (i.e., to keep the community safe) and justice systems answer to elected officials and to the public, the rationale and approach for any mental health practice must be clear to external and internal stakeholders for the practice to be adopted or sustained (Taxman and Belenko, 2011). Showing that a practice improves mental health outcomes (e.g., depressive symptoms) is sometimes not enough to promote its implementation in the justice system. Therefore, in designing the trial, we sought advice from prison providers and administrators about additional outcomes that would be persuasive to policy-makers, defensible to the public, and would promote wide buy-in for a mental health practice in the justice system. These outcomes included reduced suicidality, disciplinary and incident reports, and aggression/ victimization (each of these create hassle, cost, and liability risk for prisons), and increased enrollment and completion of correctional programs (Johnson et al., 2016). In addition, cost-effectiveness is an important guide for decisions by prisons about group IPT adoption.

Resource limitations provide another barrier to delivering MDD treatments in prison settings. As a result, specialized mental health professionals (e.g., MSWs and PhDs) are in

short supply in many prisons and prison treatment as usual typically consists of antidepressant medications, if anything (Baillargeon, Contreras, Grady, Black, and Murray, 2000). Therefore, after initially piloting group IPT in prisons (Johnson and Zlotnick, 2008; 2012), we explored a task-shifting approach. We created a more structured manual and trained additional prison counselors who were not mental health treatment specialists. A subsequent open trial (n = 22) showed that these counselors (i.e., re-entry workers and bachelor's level prison substance use counselors) were able to deliver group IPT adherently and competently (Johnson, Williams, and Zlotnick, 2015). The current trial included both master's-level mental health professionals and non-mental health specialists as counselors. Our overall goal was to find an inexpensive way to provide a community-standard dose of psychosocial MDD treatment (i.e., group IPT delivered by available prison providers) to determine whether improved outcomes would justify its use to policy-makers. In other words, could we make a standard psychotherapy treatment effective and low-cost enough to promote its adoption in prisons?

The purpose of this study was to evaluate the effectiveness and cost-effectiveness of IPT for prisoners with MDD using a rigorous RCT design. Group IPT delivered by prison providers was compared to prison mental health treatment as usual (TAU) in a sample of 181 male and female sentenced prisoners with MDD. Effectiveness outcomes, included: (1) depressive symptoms, (2) suicidality, and (3) in-prison functioning (i.e., enrollment and completion of correctional programs; disciplinary and incident reports; aggression/victimization; social support). Exploratory outcomes included time to remission from the major depressive episode, posttraumatic stress disorder symptoms, and anxiety. The study also examined social support as a mediator of the effects of IPT on depressive symptoms. By capturing the cost of group IPT for MDD, the study was able to analyze the intervention's cost-effectiveness. Finally, this trial is one of a few RCTs of correctional mental health treatments to explore gender effects, despite recognition that clinical pictures of men and women prisoners often differ (James and Glaze, 2006). Therefore, this study will explore gender as a moderator of the effectiveness of IPT relative to TAU.

Method

The trial was approved by Brown University's Institutional Review Board (FWA 00004460) and regulatory bodies overseeing prison research in participating states. A three-member external Data Safety and Monitoring Board oversaw data and safety of study participants. The trial was registered at clinicaltrials.gov (NCT01685294). Study procedures are described in detail in Johnson et al. (2016). The larger study was a Hybrid Type I implementation trial (Curran et al., 2012), a study design that tests effects of a clinical intervention on relevant outcomes (i.e., uses a rigorous RCT design to answer effectiveness questions) while observing and gathering information (i.e., pilot data) on implementation. The current manuscript reports results of the effectiveness/cost-effectiveness RCT.

Procedures

Study participants were recruited from women's facilities and from men's medium security facilities in two northeastern U.S. states. State prisons in these states serve as both prisons

and jails; this study recruited only sentenced individuals. Participating facilities (n = 6) housed the entire female state prison populations for both states. They also housed all the medium security men in one state, and part of the medium security men in the other state. Prisoner demographics and sentencing practices were similar between the two states. Treatment as usual for MDD was also similar between the two states and similar to practices at other prisons nationally (Baillargeon et al., 2000), consisting mainly of medications with occasional psychosocial treatment.

Participants and recruitment

Participants were sentenced prisoners between the ages of 18 and 65 incarcerated in one of the participating prisons. Participants also: (1) met DSM-IV-TR criteria for current primary (non- substance-induced) MDD after at least 4 weeks incarceration; and (2) reported being likely to stay at their current facilities for at least 6 months, giving us time to complete the treatment and follow-up phases. Individuals who met lifetime criteria for bipolar disorder or a psychotic disorder other than psychotic MDD were excluded, as were individuals in prison residential mental health treatment (because individuals in prison residential mental health treatment were already receiving a higher level of mental health care than that offered by IPT). Individuals who could not understand English well enough to understand the consent form or assessment instruments when read aloud were excluded. We required there to be at least 4 eligible participants at a time in a facility to randomize a new group; a few otherwise eligible participants were not at facilities at times that there were enough participants to randomize a new group (see Figure 1).

Participants were recruited through announcements made on prison housing units and through flyers. When study research staff made announcements on housing units, they handed out slips of paper to all attendees to confidentially indicate "yes" or "no" interest in finding out more about the study and then collected all the slips. Participants who found out about the study through flyers or word of mouth inside the prison submitted an interest slip to a designated prison staff person, and then study research staff collected the interest slips and met with the participant privately to explain more about the study. Screening questions included: "Is this something you might be interested in?", "Do you expect to be at this facility for at least 6 months longer?", "Have you been incarcerated for at least 4 weeks?", and the MDD screening items from the Structured Clinical Interview for DSM-IV (SCID; First, Spitzer, Gibbon, and Williams, 1996). If answers indicated eligibility, formal informed consent took place. After informed consent forms were signed, full eligibility was assessed using the SCID mood disorder module and psychotic screener sections. If the participant met criteria for the study, the rest of the baseline assessments were completed; randomization took place after the completion of the baseline assessment.

Research staff worked to minimize risk of coercion by emphasizing that participation was voluntary and had no consequences for prison treatment or terms or length of confinement. Participants received a copy of the signed consent form. Consent forms were always read aloud, and we also offered to read each questionnaire aloud.

Interventions

TAU.—Because our goal was to design a study that was relevant to prison policy decisions, we employed a treatment as usual (TAU) control condition. In order to determine the naturalistic effects and costs of adding IPT to the prison setting, participants in both conditions could receive any other treatment available to them; we did not exclude participants receiving other treatment, including antidepressant medications. We characterized TAU separately for each condition.

Participants in the TAU condition were offered referrals to prison mental health staff for TAU. Nationally, TAU for MDD within prisons typically consists of antidepressant medications (either tricyclics or SSRIs; Baillargeon et al., 2000), with about 80% of inmates being treated for MDD in one large state being given antidepressants (Baillargeon, Black, Contreras, Grady, and Pulvino, 2001). Compliance with these medications is reasonable (75–80%; Baillargeon et al., 2000). IPT was not available in participating facilities other than through our study.

IPT + TAU participants received standard group IPT, as specified in the study treatment and training manuals (Johnson and Zlotnick, 2012; Johnson et al., 2015) in addition to TAU. The therapeutic stance of IPT is active, goal-oriented, semi-structured, supportive, positive, present-focused, and conducive to skills acquisition. In this study, IPT was delivered using 20 90-minute group therapy sessions over 10 weeks with 4 individual (pre-group, mid-group, post-group, and maintenance) sessions. The maintenance session occurred approximately 4 weeks after the post-group session. The individual sessions were used successfully in our previous studies (Johnson and Zlotnick, 2008; 2012; Johnson et al., 2015) to prepare patients to use the group effectively and to keep group members focused on their interpersonal goals.

Randomization

The randomization sequence, which was blocked in varying block sizes ranging from 2 to 8, was generated at the beginning of the study by the study statistician. Condition assignments were concealed in sequentially numbered opaque envelopes, which were filled and sealed by someone not associated with the study. Study research assistants (RAs) took the envelopes to the prisons, conducted study assessments, and then opened the envelopes in the presence of study participants at the time of randomization. To increase balance between conditions, randomization was stratified by state and by sex (male or female). Because we were conducting a group treatment study, we recruited sequentially (over 2–3 weeks each) at each participating prison facility. A new IPT group began at the facility when we had randomized enough participants at a facility to begin a group; this occurred 23 times. Thus, the trial delivered IPT in 23 groups between March 3, 2012 and March 4, 2014. The mean number of participants assigned to each IPT group was 4 (range 1 – 7; median = 4). Baseline and subsequent assessments for participants in both conditions at the facility took place during the same period of time. Blinded RAs (i.e., RAs other than the randomization RA) conducted post-baseline study assessments.

Counselors

IPT counselors in the current study were employed at participating prisons, with a bachelor's degree and at least one year of experience working with incarcerated individuals. Of the 9 study counselors, 5 were employed as prison mental health clinicians (i.e., they were mental health specialty providers), and 4 were not. The 5 prison mental health clinicians were all master's-level clinicians (MA in psychology or MSW). The other 4 study counselors (i.e, the non-specialist counselors) held positions within the prison including discharge planner and public health education specialist, but were not part of the prison mental health infrastructure. Each of the 6 participating facilities had 1–3 different study counselors run groups at that facility over the course of the study (a few counselors rotated among facilities depending on their schedules).

The first author, a certified IPT supervisor, trained counselors in IPT using a 1.5-day training consisting of reviewing the treatment rationale, concepts, materials, and strategies; audio demonstrations; and live practice sessions along with feedback. The first and seventh authors provided supervision, which consisted of weekly review of therapists' audiotaped sessions and weekly individual phone consultation, with group refresher trainings twice per year.

Because this trial examined real-world effectiveness and implementation, because it used a group intervention, and because the trial was complex and the timeline tight (3 years), it was not feasible to have study counselors complete IPT practice groups before treating study participants. Some data suggest that counselors novice to IPT improve significantly after their first IPT case (Stewart et al., 2014). In a group treatment study, the 20 sessions with a counselor's first group would be considered the first "case." Therefore, we provided secondary analyses with the subsample of n = 126 IPT and TAU participants from randomization waves in which IPT participants were assigned to counselors who had led at least one previous set of IPT groups (which included the first group for two of the non-specialist counselors who had been trained in IPT for a previous study, and the second and subsequent groups for the other study counselors). This decision, described in the protocol paper, was made after the beginning of the trial but before analysis of the results (Johnson et al., 2016).

Assessments

Participants were assessed three times in prison: at baseline, at the end of group (a mean 4.2 months later), and then at approximately 3 months after the end of group (a mean of 7.0 months after baseline). Assessments included structured interviews and self-report measures, and were conducted by RAs trained in interviewer administered instruments at Brown's Clinical Assessment and Training Unit. The first author and a clinical interviewing trainer supervised the RAs. Reliability checks were made periodically throughout the study and ongoing training provided. RAs conducting post-treatment and follow-up interviews were blind to treatment assignment.

Diagnosis and screening.—Participants were asked for demographic information, trauma history (using subscales from the Trauma History Questionnaire; Hooper, Stockton, Krupnick, and Green, 2011), and arrest and sentencing history. The mood disorder module

and psychotic screener from the SCID-IV (First et al., 1996) was used to determine diagnostic eligibility.

During the follow-up period, the Longitudinal Interval Follow-up Examination (LIFE; Keller et al., 1987), a brief standardized interview, was used to assess MDD diagnosis. Unlike the SCID, which provides only a cross-sectional measure, the LIFE tracks MDD severity and course over time. The LIFE uses Psychiatric Status Ratings (PSRs) to measure MDD severity each week on a scale of 1 (asymptomatic) to 6 (incapacitated). MDD remission was defined as the first week out of episode (i.e., at a depression PSR of 1 or 2; Keller, 2003). MDD recovery was defined as a depression PSR of 1 or 2 for 8 consecutive weeks (Keller, 2003). Time to MDD remission and recovery were exploratory outcomes in the current study.

Depressive symptoms (primary outcome) was assessed with the interviewer-administered Modified 17-item Hamilton Rating Scale for Depression (HRSD; Hamilton, 1980). HRSD interrater reliability was excellent, with an intraclass correlation coefficient of .97 across raters.

Suicidality (a secondary outcome) was assessed with the Beck Scale for Suicide Ideation (BSS; Beck, Kovacs, and Weissman, 1979) which assesses current suicide ideation. However, our experience has been that many individuals in prison under-report current suicide ideation because suicide watch in prison is unpleasant. Therefore, we assessed weeks of suicide ideation during each follow-up period using the LIFE (past suicide ideation does not require mandatory reporting to the prisons), and hopelessness (using the Beck Hopelessness Scale; Beck, 1988), which is strongly related to suicidality. Suicide ideation on the LIFE is defined as a suicide PSR of "3 = occasional thoughts of suicide" or higher.

In-prison functioning.—Number of correctional programs begun and completed, number of disciplinary and incident reports, and weeks spent on punishment/isolation from baseline through follow-up were assessed through self-report, using the LIFE (Keller et al., 1987) and Timeline Followback-based (Sobell, Brown, Leo, and Sobell, 1996) calendar methods; data was reported by study week. Baseline assessments measured the 6 months prior to baseline. For post-baseline assessments, we created indices by summing similar concepts (e.g., number of correctional programs begun and completed) across available post-baseline assessment data (e.g., post-group and follow-up). If the post-baseline follow-up period exceeded 32 weeks, we truncated the data at 32 weeks to standardize the time frames for all three dependent variables (i.e., number of programs begun and completed, number of discipline and incident reports, and number of weeks spent on punishment/isolation). In our analysis, we used available data (i.e., for people still in prison at the 3-month follow-up, we added the post-treatment and 3-month follow-up counts together; for people released before the 3-month follow-up, we used the post-treatment count). Previous studies have indicated that calendar-based prisoner self-report of criminal activity and discipline problems is reliable (Sutton, Bellair, Kowalski, and Hutcherson, 2011); our experience is that it is sometimes more accurate than official correctional records. More programs begun and completed, and fewer disciplinary and incident reports would reflect better in-prison functioning.

In-prison aggression and victimization were measured using the Conflict Tactics Scale 2 (CTS2; Straus, Hamby, McCoy, and Sugarman, 1996) with adaptations to ask about incidents in relationships with anyone over the past 3 months, rather than only with romantic partners. We also removed the 7 sexual coercion subscale items because of the sensitive nature of these questions in prison. The final 32-item scale assessed CTS2 subscales of Negotiation, Psychological aggression, Physical assault, and Injury perpetration/ victimization in any relationship over the past 3 months.

Social support (an indicator of in-prison functioning, as well as a hypothesized mediator of the effects of IPT on depressive symptoms) was assessed using the Multidimensional Scale of Perceived Social Support (MSPSS; Zimet, Dahlem, Zimet, and Farley, 1988) and the UCLA Loneliness Scale (Russell, 1996).

Exploratory outcomes included an additional measure of depressive symptoms (the Quick Inventory of Depressive Symptoms – Self-Report; QIDS; Rush et al., 2003), posttraumatic stress disorder symptoms assessed using the PTSD Checklist (LEC-PCL; Ruggiero, Del Ben, Scotti, and Rabalais, 2003), and the Generalized Anxiety Disorder 7-item scale (GAD-7; Spitzer, Kroenke, Williams, and Lowe, 2006). These scales are well-established, reliable, and valid.

Moderators included SCID-I assessed number of past depressive episodes, SCID-II (First, Gibbon, Spitzer, Williams, and Benjamin, 1996) assessed Borderline and Antisocial personality disorders (BPD and ASPD), and verbal ability assessed as the sum of the correct answers to the 5 screening items from the Expressive Vocabulary Test 2 (for a total score of 0–5; Williams, 2007).

Characterization of TAU for both conditions.—We used the LIFE calendar to characterize medications taken and doses, number of individual or group counseling sessions for substance use, number of individual sessions for mental health or personal problems (including counseling or medication adjustment visits), number of group sessions for mental health or personal problems (this included prison mental health, domestic violence classes, sex offender classes, etc.), and time spent in prison residential (typically substance use) treatment programs weekly over the 6 months prior to baseline and at each follow-up period. If the post-baseline follow-up period exceeded 32 weeks, we truncated the data at 32 weeks to standardize time frames. We used post-group and follow-up assessment data for participants who completed both interviews, and post-group assessment data only for those released from prison before the 3-month follow-up.

Client satisfaction was measured using the well-validated 8-item Client Satisfaction Questionnaire-Revised (CSQ-8-R; Larsen, Attkisson, Hargreaves, & Nguyen, 1979), administered at the post-treatment assessment. Client satisfaction was assessed for IPT and for TAU separately for participants receiving both.

IPT adherence and competence.—Two licensed, Ph.D.-level experts in IPT who were independent from the study provided adherence and competence ratings by listening to audio recorded sessions using a 75-item IPT quality scale used in previous IPT studies in the

prison (Johnson and Zlotnick, 2012). Raters rated several cases together, discussed them at an item-by-item level, and then rated the rest of the tapes independently, with periodic check-ins for discussion. Overall, 96 of the 436 recorded group IPT sessions (22%) and 28 of the 246 recorded individual (i.e., pregroup, midgroup, post-group, and booster) IPT sessions (11%) were rated. For each session, three total scores were calculated: percent time spent on model (adherence), and two measures of competence: an average score across all items and an item reflecting overall quality of the session (each rated on a scale from 1 = poor, 4 = moderate, to 7 = ideal). Single-rater interrater reliability for the competence measures was moderate (Koo and Li, 2016), with an intraclass correlation coefficient of .62 across raters.

Cost.—Intervention cost was tracked by multiplying hours spent times cost per hour. Hours spent were tracked biweekly on timesheets broken down by activity (e.g., training, supervision, clinical delivery, and case notes for study counselors; listening to recordings, training counselors, and direct supervision time for study supervisors). The hourly salary of each provider was recorded. The prisons supplied information on its fringe benefit and overhead rates. Overhead primarily covers office/meeting space, equipment, administration, supplies, and local travel. Clinical supervision was costed at the local prevailing rate of $100 per hour.

Analyses

Primary analyses were *intent-to-treat.* A priori, tests of main effects were two-sided with $p = 0.05$. A priori, tests of moderation used $p < .01$ to control for Type I error. IPT+TAU and TAU alone groups were compared on values of key variables (listed in Table 1); variables with statistically significant differences ($p < .05$) were added as covariates in subsequent analyses.[1]

Missing data.—We compared treatment conditions on rates of missingness and time to missingness, and tested whether baseline characteristics (female, shorter total time incarcerated, no life sentence, state, racial/ethnic minority [yes/no], age, prior depression treatment, and number of past prison discipline/incident reports), were associated with missingness. We then employed multiple imputation (using tested variables found to be associated with missingness) to test sensitivity of study results to missing data. This method takes into account the associations between observed variables and patterns of missingness (Allison, 2002; Little and Rubin, 1987). We imputed 100 samples using regression with predictive mean matching (Schenker and Taylor, 1996) to ensure that imputed values were within the feasible range.

Footnote 1: The a priori analysis plan (Johnson et al., 2016) did not call for adjustment by differences between conditions in baseline variables due to randomization. However, randomization resulted in the sample being significantly imbalanced on two variables (borderline personality disorder and social support) that are known to be associated with MDD and IPT outcomes. The imbalance favored the TAU condition. Our team was concerned that our a priori plan (which was very conservative) may have been too conservative. We considered the debate in the larger clinical trials literature over what to do with imbalances due to randomization (i.e., covariate adjustment leads increases in power vs. spurious results; see Kahan, Jairath, Dore, and Morris, 2014; Kraemer, 2015). Upon consultation with investigators inside and outside our team, given the clinical complexity of the population (i.e., high level of comorbidity) and given the common practice of adjusting for imbalanced baseline variables in clinical trials, we ultimately decided to report analyses including these two covariates.

Model construction.—For continuous assessments, we constructed a hierarchical linear model (HLM) model to assess the difference in assessments after treatment (post group) and follow-up time periods. All HLM models included the baseline assessment value (e.g., baseline HRSD score for analyses of effects of IPT on post-baseline HRSD scores) as a covariate, a time variable (i.e., post group or follow-up assessment), treatment condition, and the condition by time interaction. We also calculated the model adjusted Cohen's d effect size estimate. For count assessments, we constructed variables reflecting the total count (e.g., of prison programs started/completed) over the entire post-baseline period (including during treatment and during the 3-month post-treatment follow-up period), and analyzed these dependent variables using zero-inflated negative binomial (ZINB) models. The ZINB models contained baseline assessment value covariates for each dependent variable index.

Preliminary analyses examined the distributions of our key variables. Preliminary analyses also explored site effects on treatment response by testing the state by condition and the sex by condition interaction in primary HLM analyses (given that correctional policies can vary between states and between men's and women's facilities). Neither state nor sex moderated the effects of IPT on the primary outcome (HRSD scores), so site effects were dropped from the analysis. We also examined changes over time within each treatment condition by evaluating time effects in the IPT and TAU samples separately. Tests of changes over time within condition and tests of moderation used $p < .01$; tests comparing treatment conditions used $p < .05$.

Analysis of the primary outcome.—We tested the hypothesis that, relative to TAU alone, IPT + TAU will result in lower depressive symptoms (HRSD scores) at post-treatment and follow-up, using hierarchical linear modeling (HLM).

Analysis of secondary outcomes.—We tested the hypotheses that, relative to TAU alone, IPT+TAU would result in lower levels of suicidality (BSS scores, Beck Hopelessness Scale scores, weeks of suicide ideation) and better in-prison functioning, including separate tests for enrollment in more correctional programs, completion of more correctional programs, fewer disciplinary and incident reports, lower levels of aggression (CTS2 aggression scores), lower levels of victimization (CTS2 victimization scores), and higher social support (MSPSS scores)/less loneliness (UCLA Loneliness Scale scores). For normally distributed variables, analyses used HLM. For count data, analyses used ZINB models.

Exploratory.—We explored whether conditions differed in time to remission and time to recovery from a depressive episode on the LIFE using Cox regression, with length of current depressive episode (in months, truncated at 120 months) and a variable reflecting yes/no lifetime "too many depressive episodes to count" as covariates. Too few participants reported active suicide ideation on the LIFE during the follow-up period to make comparisons meaningful, so we reported these results descriptively. We also explored effects of IPT on an additional measure of depressive symptoms (QIDS-SR scores), on symptoms of posttraumatic stress disorder (PTSD; LEC-PCL scores), and on anxiety symptoms (GAD-7 scores) using HLM.

Mediation analysis.—We postulated that social support, as measured by the MSPSS, would mediate the effects of IPT treatment on depressive symptoms (HRSD scores) and hopelessness (BHS scores). We used fully prospective data in testing these mediation hypotheses as recommended by Kazdin and Nock (2003): treatment assignment was determined at baseline, the social support mediator was measured at post-treatment, and the outcome measure was taken at the final outcome point. We tested for mediation using MacKinnon's product of coefficients method (MacKinnon, Lockwood, Hoffman, West, and Sheets, 2002). These mediation analyses were done both for the whole sample, and for the n = 126 subsample.

Moderator analyses.—After completion of an optimal model presenting HRSD changes from post-group phase, we tested moderators of the effect of IPT on HRSD scores (within the HLM model). Planned tests of moderators of treatment effects included sex, minority status (i.e., either Hispanic or non-White), self-reported number of lifetime arrests, life sentence (yes or no), MDD severity (i.e., baseline HRSD score), number of past depressive episodes, receiving psychotropic medications at baseline, Borderline and Antisocial personality disorders, baseline social support (MSPSS score), and verbal ability (EVT-5 screener items), with $p < .01$ to control for Type I error. Our modifier analysis strategy centered each potential modifier at a mean (or median) value for continuous values and at zero for binary values and separately tested the significance of each multiplicative interaction term (i.e., interaction between modifier and treatment condition) in predicting HRSD outcomes.

Client satisfaction was examined descriptively and t-tests were used to compare client satisfaction with IPT to client satisfaction with TAU for participants receiving both.

Cost-effectiveness was evaluated using three measures: IPT cost per additional person in remission from MDD by week 32, IPT cost per person-week in remission, and IPT cost per quality-adjusted life year (QALY) saved. QALYs saved were based on concordant findings from Jia and Lubetkin (2017) and Revicki and Wood (1998) that MDD reduced QALY levels by 56 percentage points (a 65% decline from a no-MDD QALY level of 86% to 30%).

Power.—This study was powered to detect an effect size at the lower end of the range of effect sizes of other IPT studies, or a $d = .37$ (see Johnson et al., 2016). Based on past studies (e.g., Johnson and Zlotnick, 2012), we planned for a 10% attrition rate.

Results

The trial randomized 181 participants; 91 to IPT + TAU and 90 to TAU alone. Of the 181 participants, 111 were recruited from one state (with 44 and 24 participants recruited in 5 and 3 waves respectively from the 2 participating men's facilities, and 36 and 7 participants recruited in 4 and 1 waves respectively from the 2 participating women's facilities) and 70 (49 in 7 waves from the men's facility and 21 in 3 waves from the women's facility) from the other state. Table 1 shows other baseline characteristics of the sample. Conditions did not differ significantly by any baseline characteristics except for rates of borderline personality disorder (46% of IPT + TAU group vs. 31% of TAU alone group; $p = .035$), and baseline

social support (MSPSS) scores, which were lower in the IPT + TAU group than in the TAU alone group (M = 49.8 vs. 55.9, p = .034). Therefore, these two variables were added as covariates in subsequent analyses. The CONSORT Diagram (Figure 1) shows the flow of participants through the study. Follow-up rates were 95% (n = l72) at the post-group assessment, and 86% (n = 155) at the 3-month follow-up assessment, with 96% of the sample (n = 173) providing at least one post-baseline assessment. Results reported reflect the full sample unless otherwise specified.

Preliminary analyses.

There were no significant differences between conditions on rates of missing data. Of tested variables (i.e., female, shorter total time incarcerated, no life sentence, state, racial/ethnic minority [yes/no], age, prior depression treatment, and number of past prison discipline/ incident reports), three baseline characteristics were significantly related to missingness (female, shorter total time incarcerated, no life sentence). This makes sense because the primary source of missing data was early prison release and these variables are related to shorter prison sentences. Sensitivity analyses indicated that results of multiple imputation analyses of key outcomes based on these three variables were consistent with results of the complete-data analyses. Therefore, missingness is unlikely to have affected results and we report complete-data analyses below.

Characteristics of TAU received in the 7 months between randomization and follow-up are shown in Figure 1. Results showed no evidence that participants randomized to IPT used less prison TAU over the study follow-up period. ZINB models comparing number of individual mental health sessions (counseling or medication adjustment visits), number of individual or group substance use counseling sessions, and weeks of prison residential treatment (typically for substance use) between conditions, covarying baseline values, showed no differences. Participants randomized to IPT actually attended slightly more group mental health visits (including domestic violence classes, etc.) over the 7 months after baseline (14.8 vs. 12.1), among those with non-zero responses (p = .043). There were no differences in rates of psychoactive medication use between conditions at baseline or follow-up (23% in the sample as a whole at baseline). Few people in either condition added medications or increased doses during the follow-up period (n = 4 in each condition); some dropped medications or decreased doses (n = 18 in IPT and n = 15 in TAU).

Participants in both conditions showed significant (p < .01 for within-groups tests) improvement over time in depressive symptoms (both HRSD and QIDS scores), hopelessness (BHS scores), loneliness (UCLA Loneliness scale scores), generalized anxiety symptoms (GAD scores), and PTSD symptoms (LEC-PCL scores). Only IPT participants showed significant improvement over time in suicide ideation (BSS scores). Neither group showed significant (p < .01) improvement over time in social support (MSPSS scores) or aggression or victimization (adapted CTS2 scores).

Primary outcome.

IPT + TAU produced larger reductions in depressive symptoms (assessed via the HRSD) than did TAU alone in the full (n = 181) sample (p = .025) and in the subsample of 126

participants from randomization waves where IPT participants were assigned to counselors who had led at least one previous set of IPT groups (p = .011; see Table 2).

Secondary and exploratory outcomes.

IPT + TAU produced larger reductions in BHS hopelessness scores than did TAU alone in both samples ($p < .001$; see Table 2). IPT + TAU also produced larger reductions in QIDS depressive symptom scores and in LEC-PCL posttraumatic stress disorder symptom scores than did TAU alone in both samples (see Table 2). BSS suicide ideation, CTS aggression and victimization, MSPSS social support, UCLA loneliness scores, and GAD anxiety scores did not differ significantly between conditions in either sample. Effect sizes for significant results were small to medium (see Table 2). Analyses of count outcomes (correctional programs started and completed, discipline and incident reports, and weeks spent on punishment or isolation) did not show differences between treatment conditions (see Table 3).

Figure 2 shows the growth in MDD remission over time. Controlling for length of current depressive episode and yes/no to too many past depressive episodes to count, Cox regression results confirmed that the IPT condition participants were significantly more likely to remit from their current depressive episode during the follow-up period. The hazard ratio in the full sample was 1.75, indicating that the participants in the IPT condition had a 75% higher incidence of remission during the follow-up period than did TAU participants (see Table 4). When analyses were restricted to the n = 126 subsample, IPT participants had a 105% higher incidence of remission than their TAU counterparts. Number needed to treat is a measure of clinical significance, and refers to the average number of patients who need to be treated for one of them to benefit compared with the control condition. Numbers needed to treat for MDD remission in this study were 4.86 and 4.21, respectively. Over the 32 weeks of tracking, those receiving IPT + TAU experienced 212 more person-weeks in remission than those receiving TAU alone (808 person-weeks in IPT and 596 in TAU in absolute, non-covariate adjusted numbers). That gain was entirely in the n = 126 sample assigned to providers running their second or subsequent IPT group (574 person-weeks vs. 361). The cohort attending IPT with first time IPT therapists and their controls each experienced 234 and 235 person weeks in remission, respectively.

Rates of MDD recovery (i.e., achieving remission for 8+ consecutive weeks) did not differ significantly between conditions (see Table 4). Very few participants (10 in IPT and 7 in TAU at baseline; 7 in IPT and 5 in TAU during the follow-up period) reported any weeks of active suicide ideation when assessed with the LIFE (PSR 3+). These numbers were too small for comparisons of weeks of suicide ideation between conditions to be meaningful.

Mediation.

Mediation testing failed to find any statistically significant mediation of the effects of treatment by social support either in the full sample or the n=126 subsample. Changes in social support (MSPSS) scores were significantly associated with changes in depressive symptom (HRSD) and hopelessness (BHS) scores. However, mediation was not established because treatment condition did not predict changes in social support.

Moderation.

Sex, minority status, number of arrests, having a life sentence, baseline depressive symptom severity (i.e., HRSD score), number of past depressive episodes too many to count (yes/no), being on psychotropic medications at baseline, baseline social support (MSPSS scores), and verbal ability (assessed using the EVT screener items) did not moderate the effects of IPT on depressive symptoms (i.e., HRSD scores) in the full (n = 181) sample. Borderline Personality Disorder and Antisocial Personality Disorder bordered on significant moderation (p = 0.015 and p = .015, respectively), such that the benefits of IPT relative to TAU trended toward being larger for individuals without these personality disorders.

Client satisfaction.

Client satisfaction with IPT (n = 78 forms completed after answering "yes" to "Did you receive study groups for depression?") was high, with an average score of 27.5 (SD = 5.4) on an 8 to 32-point scale. This corresponds to an item-level score of 3.43, between "3 = mostly satisfied" and "4 = very satisfied." Client satisfaction with prison TAU for depression (n = 98 total participants who answered "yes" to "Did you receive any other treatment from the prison for depression, including counseling or medications?"[2]) was 20.7 (SD = 6.7). This corresponds to an item-level score of 2.68, between "2 = indifferent or mildly satisfied" and "3 = mostly satisfied." Among clients randomized to IPT who were also receiving TAU (n = 50), satisfaction with IPT (M = 27.4, SD = 5.6) was significantly higher than was satisfaction with TAU (M = 21.4, SD = 7.5; t = 4.5, p < .001).

Adherence and competence.

Study counselors were adherent. They spent an average of 96% of time in sessions on IPT-appropriate topics and activities and 4% of session time off-model. They were reasonably competent. Scores for the overall quality item for rated sessions averaged 5.11 on a scale from 1 to 7 (SD = 1.08; range 2 – 7). Average scores across all items in rated sessions ranged from 2.74 to 6.72 (M = 4.97; SD = 0.78).

Serious adverse events during the study period included 7 medical hospitalizations (6 in IPT and 1 in TAU), 5 incidences of mental health hospitalization or being placed on prison suicide watch (3 in IPT and 2 in TAU), and one suicide attempt (in IPT). None were determined to be related or likely related to study participation.

Costs totaled $184,869. Provider costs included $47,534 for IPT provider time (i.e., for training, supervision, clinical delivery, and case notes), $13,073 in fringe benefits, and $17,736 in overhead. Supervisor consulting costs (at $100/hour) were $106,526. That equates to $2054 per person assigned to IPT. All but one of the IPT providers (including both masters'-level prison mental health counselors and those without formal mental health treatment training) and earned between $24 and $27 per hour, with one provider at $35 per hour. Of the costs, 20% was for IPT delivery (i.e., provider direct clinical hours), 8% for provider time for case notes, 60% for the interactive supervision process (provider and

---

Footnote 2: The satisfaction questionnaires asked about satisfaction with prison TAU specifically for depression; 98 participants said they were receiving non-study prison treatment for depression. Prison mental health TAU in Figure 1 refers to prison mental health and substance use care for any condition; more than 98 participants received this care.

supervisors hours spent in supervision, as well as supervisor time listening to recordings to prepare for supervision), and 12% for provider training (i.e., supervisor time providing training and counselor time receiving training). Therefore, for established IPT programs that no longer need ongoing IPT supervision (i.e., excluding all training and supervision costs), the cost per patient would be $575.

Since the IPT + TAU condition gained 212 weeks in remission relative to the TAU condition group, the cost additional person in remission was $14,211, the cost per week in remission (non-covariate adjusted, including all costs) was $758, and the cost per QALY saved was $70,354. Restricting the analysis to the more realistic scenario where providers are not running their first IPT group (i.e., the n = 126 subsample), the cost per added person in remission was $9,958, the cost per week in remission $524, and the cost per QALY $48,687. Those cost-effectiveness ratios would improve if the increased remissions continued after the week 32 data collection. For example, if the 12-person differential in remission in weeks 12–32 in the experienced provider cohort persisted for another 3 months, the cost per QALY would drop to $28,904. For established IPT programs that no longer need ongoing IPT supervision (i.e., excluding all training and supervision costs), with the 32-week follow-up period, the cost per added person in remission would be $2,788, the cost per week in remission $148, and the cost per QALY $13,632.

Discussion

MDD is a common and disabling condition that can have serious consequences for prisoners, including risk for suicide, dropout from correctional treatment programs, victimization and rejection by other inmates, aggressive acting out, and return to correctional custody. This study is the first fully-powered study of any treatment for MDD in an incarcerated population.

This goal of this randomized effectiveness study was to determine whether it was possible to provide a community-standard dose of psychosocial MDD treatment (i.e., group IPT) using available (i.e, masters-level and nonspecialist) prison counselors that was effective and low-cost enough to promote its adoption in prisons. In this effectiveness study, IPT delivered by real-world counselors improved outcomes over TAU alone. Of the 9 study counselors, 5 were master's-level prison mental health clinicians, and 4 held non-mental health positions within the prisons (e.g., discharge planner, public health specialist). Most (7) had never been trained in or used IPT before. The 2 counselors who had used IPT before held non-mental health positions within the prisons and had been trained in IPT for previous studies by our team (Johnson and Zlotnick, 2012; Johnson et al., 2015). However, these 2 counselors and one other counselor had *no formal mental health treatment training other than that provided by our studies*. Similar to Bolton et al. (2003), our previous work (Johnson et al., 2015) has found that these counselors could deliver IPT adherently and competently if provided with close supervision. This is important because trained mental health staff are in short supply in many prisons and jails. Our current and past studies have demonstrated that using non-mental health staff to provide IPT in prisons is possible. Future analyses will determine whether using non-mental health staff to provide IPT makes sense from a cost perspective given the unexpectedly low salaries of the prison mental health counselors, the fact that their

salaries were similar to those holding non mental health positions in the prisons, and the high cost of supervision (i.e., the majority of IPT costs) we observed.

The sample in this effectiveness study was also a real-world sample, with minimal exclusion criteria. The average Hamilton Depression Rating Scale score at baseline was in the "severe" range, despite most already receiving some kind of prison mental health treatment. The median number of past depressive episodes was "too many to count," and 6% were experiencing a psychotic depressive episode at study enrollment. A majority (72%) met criteria for antisocial personality disorder and 38% met criteria for borderline personality disorder. Many (42%) had attempted suicide in their lifetimes. Most (87%) had experience physical assault in their lifetimes and more than half (58%) had experienced sexual assault in their lifetimes. We did not assess pre-prison substance use in the sample, but a majority of incarcerated individuals typically have substance use disorders (Bronson, Stroop, Ziller, and Berzofsky; 2017; James and Glaze, 2006). Therefore, this was a multiply comorbid sample that only excluded lifetime mania/hypomania and primary psychotic disorders.

In terms of symptom outcomes, IPT resulted in significantly lower depressive symptoms, higher rates of MDD remission, and lower PTSD symptoms, relative to TAU alone (see Tables 2 and 4). We did not find differences in the effectiveness of IPT plus TAU relative to TAU alone by sex, minority status, baseline depressive symptom severity, being on psychotropic medications at baseline, number of arrests, having a life sentence, number of past depressive episodes, baseline social support, or verbal ability. Effect sizes for differences between conditions were small to medium. Differences between conditions were clinically significant, especially given that many in the sample appeared to have recurrent, treatment resistant MDD, multiple comorbidities, and complex psychosocial challenges.

In terms of functioning outcomes, relative to TAU alone, IPT plus prison TAU resulted in significantly lower hopelessness (p < .001), a strong correlate of suicidality and suicide risk, in both the full and subsamples. In trying to better understand this result, we examined item-level responses on the BHS in the IPT condition. We found the greatest positive changes for items indicating that the future was vague/uncertain and dark (relative to items such as "I might as well give up" or "there's no use trying"). After addressing the details and emotions of recent interpersonal problems (i.e., conflicts, grief, life changes), IPT helps individuals plan how to move forward. It is possible that the dual approach of addressing recent stressful challenges and then discussing roles and relationships (i.e., creating a picture of what life might be like) moving forward was particularly potent for depressed prisoners.

We did not find differences between conditions in other functioning outcomes (i.e., suicide ideation assessed directly, in-prison aggression and victimization, correctional programs started and completed, and discipline and incident reports). The lack of differences between conditions on these variables meant that direct cost-offsets from IPT (e.g., shorter prison time Only 30% of the sample had any discipline or incident reports during the study follow-up period, so this variable may have suffered from a floor effect. We also did not observe differences between conditions in expected mediators of IPT (social support and loneliness; see Table 2). Finally, we did not find evidence that IPT either increased or decreased use of other prison treatment services. There was a high overall level of use of other prison

treatment services in the sample. At baseline, 78% were in some kind of prison treatment (including 24% who were in residential treatment, typically for substance use), 68% were in prison mental health treatment and/or prison programming for personal problems (e.g., domestic violence and sex offender classes), and 23% were receiving psychotropic medications. Over the follow-up period, only 22% of the sample did not receive any kind of prison mental health or substance use services.

In scale-up, IPT to treat MDD among prisoners should cost at most $48,687 per QALY saved. Once start-up supervision ends and training and supervision is no longer needed, the cost per QALY drops to $13,620. This is well below the $100,000-$130,000 threshold commonly viewed as an acceptable maximum for safety investment in the United States (Glick et al., 2015, Miller and Hendrie, 2012).

Cost-effectiveness analyses provided insights into the costs and benefits of newer versus more established IPT programs. Effects of IPT on weeks of MDD remission (on which cost-effectiveness was based), though significant in the overall (n = 181) sample, were driven by differences observed entirely in the sample of n = 126 participants randomized in waves in which IPT participants were assigned to counselors who had led at least one previous set of IPT groups. Between-conditions effect sizes of d = 0.30 – 0.40 were also observed primarily in the n = 126 subsample (see Table 2). Therefore, similar to other IPT studies (Stewart et al., 2014), it is possible that our therapists (regardless of their level of prior mental health training) were more efficient at improving outcomes using IPT after their first experience with IPT. The study clinical supervisors' impressions were that this was the case. Therefore, we suggest continuing supervision in IPT through at least two training cases. However, given this initial investment in supervision, our results suggest that cost-effectiveness of IPT programs will improve with more established IPT programs versus newer programs because (1) counselors would have led at least one set of IPT groups, and (2) training and supervision costs, which accounted for 72% of the costs of IPT in this study, would diminish or end with time as therapists became more experienced in IPT.

Strengths and limitations

Strengths of this study include a rigorous study design (i.e., randomized trial, blinded research assistants, strong interrater reliability for interviewer-rated assessments, use of a treatment manual and fidelity ratings, recruiting the full targeted sample size) with strong generalizability and effectiveness components (minimal exclusion criteria, use of prison mental health and other available counselors) in a real-world setting. Rigorous mental health treatment studies of any kind of mental health treatment for any disorder in incarcerated populations are rare. Given the mental health burden in this population, they are also important. The primary study weakness was that the 3-year project period and group treatment design did not allow for study counselors to have training cases, necessitating secondary analysis with the subset of the sample randomized in waves in which counselors had run at least one previous IPT group. In addition, unexpectedly low rates of self-reported suicide ideation and of discipline/incident reports and time spent in isolation in the sample reduced power for these outcomes. The tight project period and short sentence lengths in the states where we were working also did not allow for follow-up beyond 3 months post-

treatment. Finally, because the study focused on during-incarceration outcomes, it did not address post-release effects of treatment (such as potential better post-release functioning, changes in recidivism, etc.).

Conclusions and recommendations

IPT is feasible, acceptable, effective, and cost-effective, and we recommend its use for MDD among prisoners. Our results suggested that adding IPT to prison TAU reduces depressive symptoms, PTSD symptoms, and hopelessness and increases MDD remission among prisoners with MDD. Previous studies (Johnson and Zlotnick, 2012; Johnson et al., 2015) had already established the feasibility and acceptability of IPT for use in prisons. Having IPT as an evidence-based treatment option is useful because: (1) it is currently the only evidence-based treatment for MDD tested in an incarcerated population; (2) prisoners may want the opportunity to address life events and interpersonal conflicts related to MDD (which is not an option with medication alone); and (3) as observed in this study, many prisoners may still be experiencing depression despite receiving other treatments. Although IPT was effective for MDD, we did not observe many of the prison functioning benefits we hypothesized for IPT (i.e., reductions in prison aggression/victimization, discipline and incident reports, time spent in punishment/isolation, use of other services). These outcomes would have provided additional implementation motivators for resource-challenged correctional settings.

Effectiveness results from this Hybrid Type I implementation trial inform clinical practice and provide a foundation for implementation of IPT for a severe and prevalent disorder among a vulnerable, understudied population. Subsequent articles will report pilot implementation data collected, including implementation barriers and facilitators in prison settings. They will also report finer-grained analysis of counselor factors, such as the relationship between fidelity and study outcomes, clinical and supervision hours required to reach fidelity for counselors with and without prior mental health treatment experience, and costs and cost-effectiveness of each kind of counselor.

Acknowledgement.

References

Allison PD (2002). Missing Data. Thousand Oaks, CA: Sage Publications.

Baillargeon J, Binswanger IA, Penn JV, Murray OW, Williams BA (2009). Psychiatric disorders and repeat incarcerations: The revolving prison door. American Journal of Psychiatry, 166(1), 103–109. [PubMed: 19047321]

Baillargeon J, Black S, Contreras S, Grady J, & Pulvino J (2001). Anti-depressant prescribing patterns for prison inmates with depressive disorders. Journal of Affective Disorders, 63, 224–231.

Baillargeon J, Contreras S, Grady JJ, Black SA, Murray O (2000). Compliance with antidepressant medication among prison inmates with depressive disorders. Psychiatric Services, 51(11), 1444–1446. [PubMed: 11058195]

Baillargeon J, Penn JV, Thomas CR, Temple JR, Baillargeon G, Murray OJ (2009). Psychiatric Disorders and Suicide in the Nation's Largest State Prison System. Journal of the American Academy of Psychiatry and the Law, 37, 188–193. [PubMed: 19535556]

Beck A (1988). Beck Hopelessness Scale: The Psychological Corporation.

Beck A, Kovacs M, Weissman A (1979). Assessment of Suicidal Ideation: The Scale for Suicidal Ideation. Journal of Consulting and Clinical Psychology, 47(2).

Blitz C, Wolff N, Shi J (2008). Physical victimization in prison: The role of mental illness. International Journal of Law and Psychiatry, 31, 385–393. [PubMed: 18809210]

Bolton P, Bass J, Neugebauer R, Verdeli H, Clougherty KF, Wickramaratne P, Speelman L, Ndogoni L, Weissman M (2003). Group interpersonal psychotherapy for depression in rural Uganda: A randomized controlled trial. Journal of the American Medical Association, 289(23), 3117–3124. [PubMed: 12813117]

Brady TM, Krebs CP, & Laird G (2004). Psychiatric comorbidity and not completing jail-based substance abuse treatment. American Journal on Addiction, 13, 83–101.

Bronson J, Stroop J, Zimmer S, & Berzofsky M (2017). Drug Use, Dependence, and Abuse Among State Prisoners and Jail Inmates, 2007–2009. Bureau of Justice Statistics (NCJ 250546). Retrieved from https://www.bjs.gov/content/pub/pdf/dudaspji0709.pdf.

Carson E (2015). Prisoners in 2014. Retrieved from www.bjs.gov/content/pub/pdf/p14.pdf.

Deboska A, & Boduszek D (2017). Child abuse and neglect profiles and their psychosocial consequences in a large sample of incarcerated males. Child Abuse & Neglect, 65, 266–277. [PubMed: 27989568]

Elkin I, Shea MT, Watkins JT, Imber SD, Sotsky SM, Collins JF, Glass DR, Pilkonis PA, Leber WR, Docherty JP, Fiester SJ, & Parloff MB (1989). National Institute of Mental Health Treatment of Depression Collaborative Research Program: General effectiveness of treatments. Archives of General Psychiatry, 46, 971–982. [PubMed: 2684085]

Fazel S, & Danesh J (2002). Serious mental disorder in 23,000 prisoners: A systematic review of 62 surveys. Lancet, 359, 545–550. [PubMed: 11867106]

Ferrari AZ, Charlson FJ, Norman RE, Patten SB, Freedman G, Murray CJL, Vos T, Whiteford HA (2013). Burden of depressive disorders by country, sex, age, and year: Findings from the Global Burden of Disease Study 2010. PLOS Medicine 10 (11): e1001547. [PubMed: 24223526]

First MB, Gibbon M, Spitzer RL, Williams JBW, & Benjamin LS (1996). Structured Clinical Interview for DSM Axis II Personality Disorders. New York: Biometrics Research Department.

First MB, Spitzer RL, Gibbon M, & Williams JBW (1996). Structured Clinical Interview for DSM-IV Axis I Disorders: Patient Edition. New York: Biometrics Research Department.

Glick HA, McElligott S, Pauly MV, Willke RJ, Bergquist H, Doshi J, Fleisher LA, Kinosian B, Perfetto E, Polsky DE, Schwartz JS (2015). Comparative effectiveness and cost-effectiveness analyses frequently agree on value. Health Affairs, 34(5), 805–811. [PubMed: 25941282]

Hamilton. (1980). Rating depressive patients. Journal of Clinical Psychiatry, 41(12, sec2), 21–24. [PubMed: 7440521]

Hiller ML, Knight K, & Simpson DD (1999). Prison-based substance abuse treatment, residential aftercare and recidivism. Addiction, 94(6), 833–842. [PubMed: 10665073]

Hooper LM, Stockton P, Krupnick JL, & Green BL (2011). Development, use, and psychometric properties of the Trauma History Questionnaire. Journal of Loss and Trauma, 16, 258–283.

James D, & Glaze LE (2006). Mental Health Problems of Prison and Jail Inmates. Bureau of Justice Statistics (NCJ 213600). Retrieved from https://www.bjs.gov/content/pub/pdf/mhppji.pdf.

Jia H, & Lubetkin EI (2017). Incremental decreases in quality-adjusted life years (QALY) associated with higher levels of depressive symptoms for U.S. Adults aged 65 years and older. Health and Quality of Life Outcomes, 15(1), 9. [PubMed: 28077154]

Johnson JE, Miller T, Stout RL, Zlotnick C, Cerbo LA, Andrade JT, & Wiltsey-Stirman S (2016). Study protocol: Hybrid Type I cost-effectiveness and implementation study of interpersonal

psychotherapy (IPT) for men and women prisoners with major depression. Contemporary Clinical Trials, 47, 266–274. [PubMed: 26845030]

Johnson JE, Williams C, Zlotnick C (2015). Development and feasibility of a cell phone-based transitional intervention for women prisoners with comorbid substance use and depression. The Prison Journal, 95(3), 330–352. [PubMed: 26508805]

Johnson JE, & Zlotnick C (2008). A pilot study of group interpersonal psychotherapy for depression in substance-abusing female prisoners. Journal of Substance Abuse Treatment, 34(4), 371–377. [PubMed: 17629441]

Johnson JE, & Zlotnick C (2012). Pilot study of treatment for major depression among women prisoners with substance use disorder. Journal of Psychiatric Research, 46(9), 1174–1183. [PubMed: 22694906]

Kahan BC, Jairath V, Dore CJ, & Morris TP (2014). The risks and rewards of covariate adjustment in randomized tirlas: Ann assessment of 12 outcomes from 8 studies. Trials, 15, 139. [PubMed: 24755011]

Kazdin AE, & Nock MK (2003). Delineating mechanisms of change in child and adolescent therapy: Methodological issues and research recommendations. Journal of Child Psychology and Psychiatry, 44(8), 1116–1129. [PubMed: 14626454]

Keaveny M, Zauszniewski JA (1999). Life events and psychological well-being in women sentenced to prison. Issues in Mental Health Nursing, 20(1), 73–89. [PubMed: 10196925]

Keller M (2003). Past, present, and future directions for defining optimal treatment outcome in depression: Remission and beyond. Journal of the American Medical Association, 289, 3152–3160. [PubMed: 12813121]

Keller M, Lavori P, Friedman B, Nielsen E, Endicott J, McDonald-Scott N, Andreasen N (1987). The longitudinal interval follow-up evaluation: A comprehensive method for assessing outcome in prospective longitudinal studies. Archives of General Psychiatry, 44, 540–548. [PubMed: 3579500]

Klein DN, & Santiago NJ (2003). Dysthymia and chronic depression: Introduction, classification, risk factors, and course. Journal of Clinical Psychology/In Session, 59(8), 807–816. [PubMed: 12858423]

Koo T, Li MY (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine, 15(2), 155–163. [PubMed: 27330520]

Kraemer HC (2015). A source of false findings in published research studies: Adjusting for covariates. JAMA Psychiatry, 72 (10), 961–962. [PubMed: 26244634]

Larsen D, Attkisson CC, Hargreaves WA, & Nguyen TD (1979). Assessment of client/patient satisfaction: Development of a general scale. Evaluation and Program Planning, 2, 197–207. [PubMed: 10245370]

Little R, & Rubin D (1987). Statistical Analysis with Missing Data. New York: Wiley and Sons.

MacKinnon D, Lockwood C, Hoffman J, West S, & Sheets V (2002). A comparison of methods to test mediation and other intervening variable effects. Psychological Methods, 7(1), 83–104. [PubMed: 11928892]

Marcus D, Hamlin RJ, Lyons PM (2001). Negative affect and interpersonal rejection among prison inmates in a therapeutic community: A social relations analysis. Journal of Abnormal Psychology, 110(4), 544–552. [PubMed: 11727944]

McDaniels-Wilson C, & Belknap J (2008). The extensive sexual violation and sexual abuse histories of incarcerated women. Violence Against Women, 14, 1090–1127. [PubMed: 18757348]

Miller TR, & Hendrie D (2012). Economic evaluation of injury prevention and control programs In Li G, Baker S (ed.), Injury Research: Theories, Methods and Approaches, New York: Springer, 641–666.

Minton T, Zeng Z (2015). Jail inmates at midyear 2014. Retrieved from http://www.bjs.gov/index.cfm?ty=pbdetail&iid=5299.

National Collaborating Centre for Mental Heath. (2010). Depression: The NICE guideline on the treatment and management of depression in adults, updated edition. Retrieved from https://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0016605/.

Parsons ML, Warner-Robbins C (2002). Factors that support women's successful transition to the community following jail/prison. Health Care for Women International, 23, 6–18. [PubMed: 11822560]

Revicki DA, & Wood M (1998). Patient-assigned health state utilities for depression-related outcomes: differences by depression severity and antidepressant medications. Journal of Affective Disorders, 48(1), 25–36. [PubMed: 9495599]

Ruggiero KJ, Del Ben K, Scotti JR, Rabalias AE (2003). Psychometric Properties of the PTSD Checklist-Civilian Version. Journal of Traumatic Stress, 16(5), 495–502. [PubMed: 14584634]

Rush A, Trivedi MH, Ibraham HM, Carmody TJ, Arnow B, Klein DN, Markowitz JC, Ninan PT, Kornstein S, Manber R, Thase ME, Kocsis JH, & Keller MB (2003). The 16-item quick inventory of depressive symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): A psychometric evaluation in patients with chronic major depression. Biological Psychiatry, 54(5), 573–583. [PubMed: 12946886]

Russell D (1996). The UCLA Loneliness Scale (Version 3): Reliability, validity, and factor structure. Journal of Personality Assessment, 66, 20–40. [PubMed: 8576833]

Skeem J, Louden JE, Manchak S, Haddad E, & Vidal S (2009). Social networks and social control of probationers with co-occurring mental and substance abuse problems. Law and Human Behavior, 33(2), 122–135. [PubMed: 18581220]

Sobell LC, Brown J, Leo GI, Sobell MB (1996). The reliability of the Alcohol Timeline Followback when administered by telephone and by computer. Drug and Alcohol Dependence, 42, 49–54. [PubMed: 8889403]

Spitzer R, Kroenke K, Williams JB, Lowe B (2006). A brief measure for assessing generalized anxiety disorder. Archives of Internal Medicine, 166, 1092–1097. [PubMed: 16717171]

Stewart MO, Raffa SD, Steele JL, Miller SA, Clougherty KF, Hinrichsen GA, & Karlin BE (2014). National dissemination of interpersonal psychotherapy for depression in veterans: Therapist and patient-level outcomes. Journal of Consulting and Clinical Psychology, 82(6), 1201–1206. [PubMed: 25045906]

Straus MA, Hamby SL, McCoy SB, & Sugarman DB (1996). The revised conflict tactic scales (CTS2). Journal of Family Issues, 17, 283–316.

Sutton J, Bellair PE, Kowalski RL, Hutcherson DT (2011). Reliability and validity of prisoner self-reports gathered using the life event calendar method. Journal of Quantitative Criminology, 27(2), 151–171. [PubMed: 24031156]

Taxman FS, & Belenko S (2011). Implementing Evidence-Based Practices in Community Corrections and Addiction Treatment. New York: Springer.

Varese T, Pelowski S, Riedel H, & Heiby EM. (1998). Assessment of cognitive-behavioral skills and depression among female prison inmates. European Journal of Psychological Assessment, 14(2), 141–145.

Weinberger A, McKee SA, & Mazure CM (2010). Inclusion of Women and Gender-Specific Analyses in Randomized Clinical Trials of Treatments for Depression. Journal of Women's Health, 19(9), 1727–1732.

Weissman MM, Markowitz JC, & Klerman GL (2000). Comprehensive guide to interpersonal psychotherapy. New York: Basic Books.

Williams K (2007). Expressive Vocabulary Test, Second Edition. Bloomington, MN: Pearson.

Zimet GD, Dahlem NW, Zimet SG, & Farley GK (1988). The Multidimensional Scale of Perceived Social Support. Journal of Personality Assessment, 52, 30–41.

Public Health Significance:

Major depressive disorder (MDD) is a common and disabling condition among the 2.3 million individuals incarcerated in the United States on any given day. Interpersonal psychotherapy (IPT) reduces depression, posttraumatic disorder symptoms, and hopelessness, and increases rates of remission from MDD among prisoners.
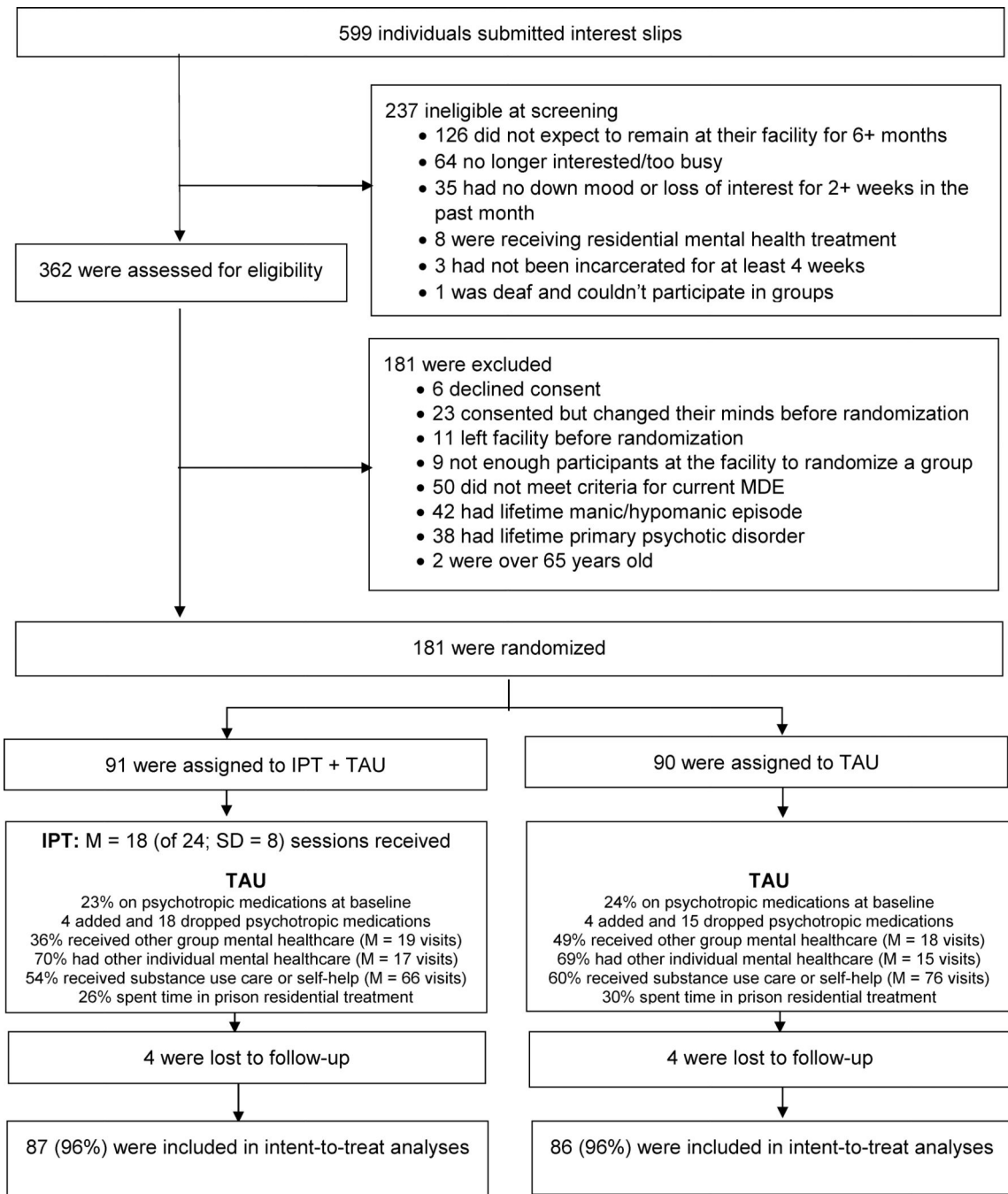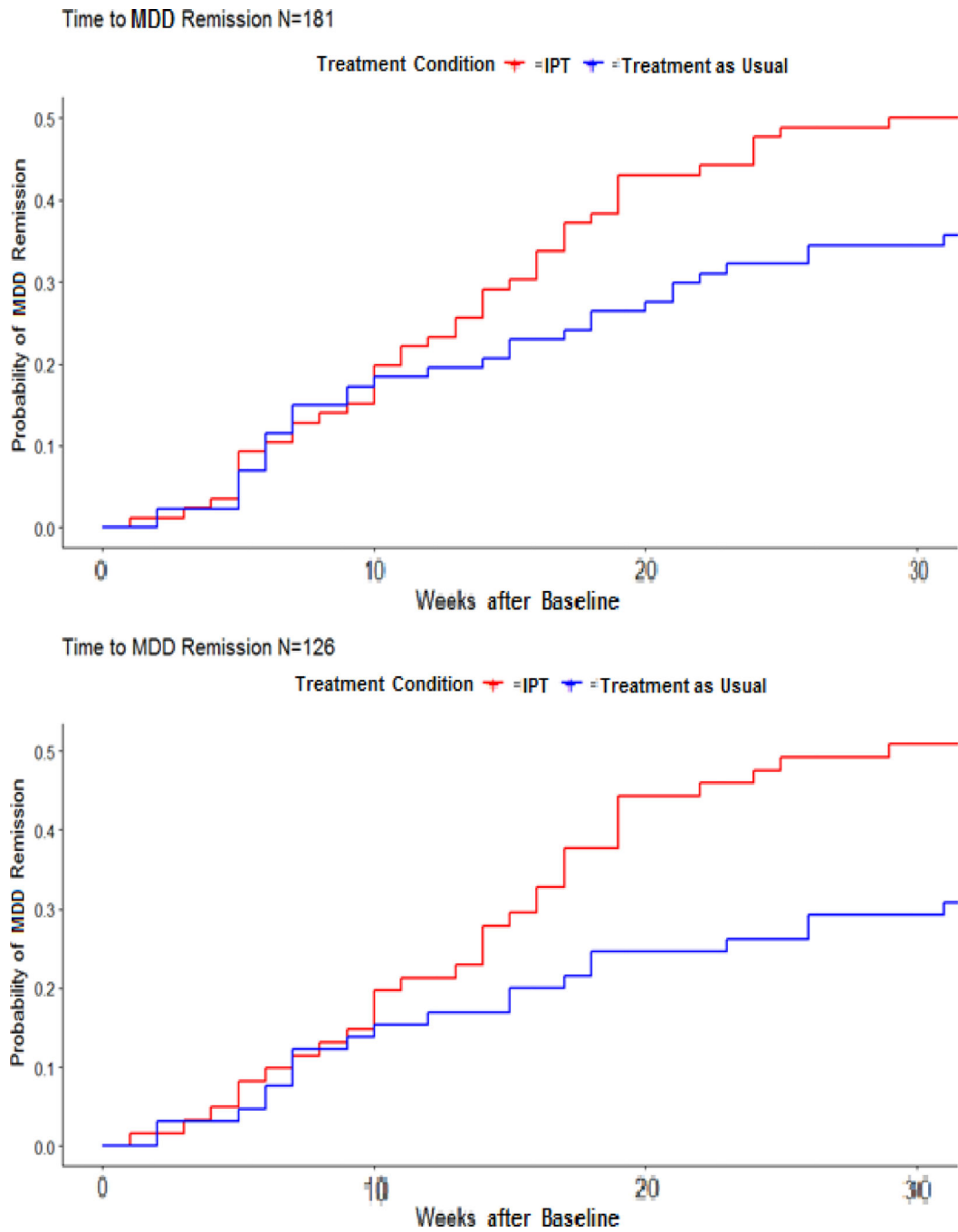
Figure 1:
Consort Diagram

Figure 2:
Time to MDD remission in the full sample (n = 181) and in the n = 126 subsample

Table 1:

Sample descriptors

	Sample (n = 181)
Age – mean	39 (SD 10.4; range 20–61)
Female – n (%)	64 (35%)
Hispanic – n (%)	34 (19%)
Race – n (%)	
African-American/Black	36 (20%)
Asian	2 (1%)
Native American/Alaskan Native	7 (4%)
White	112 (62%)
Other	24 (13%)
High school graduate or received GED – n (%)	140 (77%)
Employed prior to prison – n (%)	100 (56%)
Legal income in year prior to prison – n (%)	
Under $10,000	88 (49%)
$10,000 - $20,000	26 (14%)
$20,000 - $30,000	22 (12%)
$30,000 or more	41 (23%)
Number or prior arrests (self-report)	Median = 6 (range 0 – 100)
Time served on current sentence (in months)	Median = 31 (range 1 – 489)
Life sentence – n (%)	27 (15%)
In other MH treatment at baseline	
Group or individual MH appointments (non-IPT)	123 (68%)
Psychotropic medications	41 (23%)
Current MDE with psychotic features – n (%)	11 (6%)
Number of past depressive episodes	Median = too many to count (range 0 – too many to count)
More than 10 lifetime depressive episodes	131 (72.4%)
Attempted suicide in lifetime – n (%)	76 (42%)
Antisocial personality disorder – n (%)	131 (72%)
Borderline personality disorder – n (%)	69 (38%)
Was physically assaulted in lifetime – n (%)	158 (87%)
Was sexually assaulted in lifetime – n (%)	105 (58%)
Baseline values of outcome measures	
Hamilton Rating Scale for Depression – M (SD)	25.6 (SD 7.3)
Beck Hopelessness Scale - M (SD)	7.72 (SD 5.88)
Beck Suicide Scale - M (SD)	2.62 (SD 4.15)
Adapted CTS2 Aggression score - M (SD)	35.4 (SD 23.65)
Adapted CTS2 Victimization score - M (SD)	32.88 (SD 13.66)
Multidimensional Scale of Perceived Social Support - M (SD)	52.85 (SD 18.73)
UCLA Loneliness Scale – M (SD)	29.65 (SD 6.03)
Quick Inventory of Depressive Symptoms – M (SD)	13.05 (SD 4.18)

	Sample (n = 181)
PTSD Checklist – Civilian Version – M (SD)	45.72 (SD 12.86)
Generalized Anxiety Disorder-7 – M (SD)	12.04 (SD 5.42)

Table 2:

Continuous outcomes of IPT + TAU versus prison TAU

	n = 181					n = 126 from randomization waves where IPT participants were assigned to counselors who had run 1+ previous IPT group				
	IPT mean score[a]	TAU mean score[a]	F	p	Cohen's d	IPT mean score[a]	TAU mean score[a]	F	p	Cohen's d
Hamilton Rating Scale for Depression	14.61	16.36	5.12	**0.025**	−0.18	14.05	17.10	6.62	**0.011**	−0.31
Beck Hopelessness Scale	4.20	6.15	16.85	**<0.001**	−0.38	4.11	6.20	14.30	**<0.001**	−0.40
Beck Suicide Scale	1.37	1.72	0.19	0.661	−0.12	1.31	1.66	0.02	0.883	−0.14
Adapted CTS2 Aggression	29.08	29.90	0.73	0.393	−0.09	29.42	30.10	0.74	0.391	−0.07
Adapted CTS2 Victimization	30.26	31.51	2.60	0.109	−0.13	30.43	31.15	1.14	0.287	−0.07
Multidimensional Scale of Perceived Social Support (MSPSS)	56.25	53.94	2.18	0.141	0.13	54.88	52.35	2.19	0.142	0.15
UCLA Loneliness Scale	26.24	27.64	2.24	0.137	−0.23	26.66	27.83	1.06	0.306	−0.20
Quick Inventory of Depressive Symptoms	8.30	9.43	4.09	**0.045**	−0.21	7.98	9.82	7.06	**0.009**	−0.33
PTSD Checklist - Civilian Version	35.91	38.75	4.03	**0.047**	−0.19	34.38	39.32	5.71	**0.019**	−0.34
General Anxiety Disorder-7	7.98	8.87	1.36	0.245	−0.13	7.45	9.17	3.14	0.079	−0.27

Note: Higher MPSS scores are better; lower scores on all the other scales are better

[a]Least squares mean scores for the post-treatment and follow-up assessments

Table 3:

Count outcomes of IPT + TAU versus prison TAU from baseline through follow-up

	Did this happen?					If yes, how many times?				
	% yes	β*	SE (β)	t	p	range	β*	SE (β)	t	p
n = 181										
Correctional programs started and completed (higher is better)	87%	0.09	0.13	0.67	0.502	1–18	0.84	1.02	0.83	0.406
Discipline and incident reports (lower is better)	42%	−0.41	0.31	−1.31	0.192	1–8	−0.74	0.69	−1.07	0.285
Weeks spent on punishment or isolation (lower is better)	27%	−0.48	0.40	−1.21	0.227	3–100%	0.30	0.56	0.54	0.592
n = 126										
Correctional programs started and completed (higher is better)	89%	0.14	0.14	1.02	0.309	1–18	−0.41	1.23	−0.33	0.738
Discipline and incident reports (lower is better)	42%	−0.59	0.34	−1.73	0.084	1–8	−0.64	0.72	−0.89	0.375
Weeks spent on punishment or isolation (lower is better)	27%	−0.35	0.44	−0.81	0.417	3–100%	0.24	0.56	0.64	0.523

*
A negative sign means that the value was lower in IPT than in TAU; a positive sign means that the value was higher in IPT than in TAU.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4:

Time to MDD remission and time to MDD recovery

	Hazard ratio (95% CI)	*p*
Time to MDD remission (n = 181)	1.75 (1.09 – 2.82)	**0.021**
Time to MDD recovery (n = 181)	1.28 (0.68 – 2.41)	0.450
Time to MDD remission (n = 126)	2.05 (1.15 – 3.64)	**0.015**
Time to MDD recovery (n = 126)	1.62 (0.73 – 3.57)	0.235