Benchmarking Treatment Effectiveness of Community-Delivered Trauma-Focused Cognitive Behavioral Therapy

Brittany N. Rudd,1 Briana S. Last,2 Courtney Gregor,1 Kamilah Jackson,3 Steven Berkowitz,1,4 Arturo Zinny,3,5 Hilary E. Kratz,1,6 Lauren Cliggitt,7 Danielle R. Adams,1,8 Lucia M. Walsh,1,9 and Rinad S. Beidas1,10,11

Highlights

- First effectiveness study to benchmark TF-CBT against efficacy studies.
- Study evaluates the effectiveness of TF-CBT in urban community mental health clinics.
- Youth experienced small improvements in their PTSD symptoms, PTSD functioning, and general mental health.
- Magnitude of effect sizes were smaller than in original efficacy studies.
- There is room for improvement with regard to trauma treatments for youth in urban settings.

Abstract

It is critical for urban youth with post-traumatic stress disorder (PTSD) living in poverty to have access to evidence-based interventions for their traumatic stress. However, there is limited research on the effectiveness of these interventions when provided in urban, community settings. The objectives of the current study are to (a) evaluate the effectiveness of trauma-focused cognitive behavioral therapy delivered from 2013 to 2016 in 15 behavioral health agencies on youth (N = 114) PTSD as well as general mental health symptoms and functioning, and (b) benchmark these clinical outcomes against other published efficacy and effectiveness trials. Effectiveness data are from the Philadelphia County Community Behavioral Health System, a system that has invested significantly in the training and ongoing support of clinicians providing high-quality trauma services to youth since 2012. From baseline to last assessment, youth PTSD symptom severity (d = 0.34), PTSD functional impairment (d = 0.38), and overall mental health problem severity (d = 0.29) improved. The effect sizes of improvements were smaller than effect sizes observed in efficacy and effectiveness studies. This study is the first benchmarking study of TF-CBT and provides preliminary findings with regard to the effectiveness, and transportability, of TF-CBT to urban community settings that serve youth in poverty.

Keywords Trauma-focused cognitive behavioral therapy · Post-traumatic stress disorder · Effectiveness research · Benchmarking · Community-delivered intervention

Introduction

Exposure to traumatic events in childhood and adolescence is a serious public health concern (Gillespie et al., 2009; Overstreet & Mathews, 2011) that affects the
majority of youth in the United States (Finkelhor, Ormrod, & Turner, 2009). A recent meta-analysis found that approximately 16% of all youth who experience a traumatic event go on to develop post-traumatic stress disorder (PTSD; Alisic et al., 2014), a psychiatric disorder characterized by the experience of intrusive thoughts and memories of the traumatic event, avoidance of reminders of the traumatic event, altered cognitions and mood, and hyperarousal (APA, 2013). Due to the often severe and impairing nature of PTSD symptoms, it is vital for youth to have access to evidence-based interventions for trauma in their communities.

Trauma-focused cognitive behavioral therapy (TF-CBT) is an evidence-based treatment for youth with PTSD (Cohen, Mannarino, & Deblinger, 2016) ages 3–18. TF-CBT is a structured, short-term cognitive behavioral therapy that focuses on improving trauma-related problems, resiliency, adaptive functioning, and parent–child relationships. The treatment is component and phase-based and includes the following modules: (a) psychoeducation and parenting skills, (b) relaxation skills, (c) affective regulation skills, (d) cognitive processing skills, (e) trauma narrative and processing, (f) in vivo mastery of trauma reminders, (g) conjoint child–parent sessions, and (h) enhancing safety and social skills. In the landmark, multi-site (i.e., two university-affiliated clinics) randomized controlled efficacy trial of TF-CBT versus child-centered therapy, an intervention that attempts to improve problematic parent–child relationships, youth who received TF-CBT demonstrated significantly more improvement with regard to PTSD, depression, and behavioral problems, as well as shame and abuse-related attributions (Cohen, Deblinger, Mannarino, & Steer, 2004).

A recent meta-analysis was conducted of RCTs that compared active trauma-focused interventions (N = 39) to active or passive control conditions among children and adolescents experiencing PTSD (Morina et al., 2016). The authors used the term “trauma-focused” to refer to any intervention informed by cognitive behavioral therapy that focused on the memory of the trauma and/or its meaning through the use of cognitive therapy and/or prolonged exposure, which is thought to be the key mechanism of TF-CBT. Control conditions were subdivided into waitlist as well as active control conditions such as providing treatment as usual, supportive counseling, and psychoeducation. Trauma-focused interventions produced a large effect size when compared to waitlist control at posttreatment (g = 1.44; k = 8) and a medium to large effect size when compared to active control conditions at posttreatment (g = 0.56; k = 7) on PTSD symptoms. A large effect size was also found when trauma-focused interventions were compared to active control conditions at a longer follow-up (i.e., 3–24 months; a comparison to a waitlist control at follow-up was not possible given few trials, g = 0.66). These findings suggest the efficacy of trauma-focused interventions, like TF-CBT, as a treatment approach for PTSD.

Efficacy studies, like the majority of those included in the Morina et al. (2016) meta-analysis, seek to precisely estimate average treatment effects in tightly controlled RCTs (Chambless & Hollon, 1998). However, using research methods that optimize internal validity sacrifices the generalizability of results to community samples (i.e., external validity, see Persons & Silberschatz, 1998 for review). A large, growing body of research suggests that efficacy studies do not always generalize to the community members for whom the intervention was developed (Susukida, Crum, Ebnesajjad, Stuart, & Mojtabai, 2017).

Effectiveness studies, on the other hand, are designed to evaluate how well a treatment works in the community by optimizing external validity. Studying the effectiveness of TF-CBT delivered in the community is a critical step in addressing barriers and facilitators to its widespread use. In the only multi-site (i.e., eight community behavioral health centers) randomized controlled effectiveness study of TF-CBT versus treatment as usual in Norway, youth who received TF-CBT reported significantly lower levels of PTSD, depression, and general mental health symptoms (Jensen et al., 2014). Moreover, recent statewide non-randomized controlled effectiveness trials of TF-CBT provided in community settings in Connecticut (Donisch, 2018) and Delaware (Webb, Hayes, Grasso, Laurenceau, & Deblinger, 2014) support the effectiveness of TF-CBT. However, more research is needed on the effectiveness of TF-CBT, particularly for youth most severely affected by trauma.

Youth living in large cities in poverty, including ethnic/racial minorities, experience the highest rates of traumatic experiences compared to their peers who live in suburban or rural communities (Sabol, Coulton, & Polousky, 2004; Snyder & Sickmund, 2006), putting them at great risk for developing PTSD (Foster, Kuperminc, & Price, 2004; Gillespie et al., 2009). Further, they experience more severe PTSD symptoms and impairment, are less likely to recover, and more likely to be re-traumatized than their peers (Bonanno, Galea, Bucchiarelli, & Vlahov, 2007; Brewin, Andrews, & Valentine, 2000). Over the last two decades, a growing body of work supports the association between the experience of discrimination and poorer mental health among marginalized groups, particularly the development of trauma-like symptoms (Pascoe & Smart Richman, 2009). Marginalized youth often experience ongoing traumas, rather than a single traumatic event. While these youth have the greatest mental health need, there are large disparities in access to mental health services. In a recent nationally representative study of
children and young adults, Black and Latinx youth visit rates to outpatient mental health services were half those of non-Latinx White youth, even when controlling for mental health impairment, demographics (i.e., age, gender, and census region), and insurance status (Marrast, Himmelstein, & Woolhandler, 2016). Efforts are underway to transport evidence-based interventions like TF-CBT to youth in urban, community settings (Beidas, Adams et al., 2016). However, the question remains: once transported to urban settings with community providers, are the gold standard, efficacious evidence-based interventions effective?

The current study aims to: (a) preliminarily evaluate the effectiveness of TF-CBT delivered by community clinicians in 15 Philadelphia community behavioral health agencies from 2013 to 2016 on improving youth outcomes and benchmark youth clinical outcomes against other published efficacy and effectiveness trials. Clinical outcomes were compared with the multi-site randomized controlled efficacy trial of 229 youth who experienced sexual abuse in the United States (Cohen, Deblinger et al., 2004, N = 114 received TF-CBT), and the only multi-site randomized controlled effectiveness trial of 156 youth who experienced heterogeneous traumas in Norway (Jensen et al., 2014, N = 79 received TF-CBT). We hypothesized that (a) youth PTSD and overall mental health symptoms and functioning would improve over the course of treatment, and (b) effect sizes in this sample would be less robust than the efficacy study but similar in magnitude to the effectiveness study.

Method

Setting

Philadelphia is a large, diverse urban city of over 1.5 million residents, 25% of which are younger than 20 years old. Philadelphia’s residents are 41% African American, 35% Non-Latinx White, 14% Latinx, 7% Asian, and 3% other (PEW, 2017). Philadelphia has high rates of violent crime: 989 per 100,000 residents for 2016 (PEW, 2017) and is the poorest of the nation’s 10 largest cities, with 25% of its residents living below the poverty level and the majority of its youth enrolled in Medicaid. Public behavioral health services, financially supported by Medicaid, are managed by Community Behavioral Health (CBH), a non-profit managed care organization (i.e., “carve-out”) established by the City of Philadelphia that functions as a component of the Department of Behavioral Health and Intellectual disAbility Services (DBHIDS). DBHIDS oversees all public behavioral health service delivery in Philadelphia County.

In 2011, DBHIDS initiated a full-scale effort to develop a trauma-informed behavioral health system due to the high rates of trauma exposure in Philadelphia. In 2012, DBHIDS was awarded a National Child Traumatic Stress Initiative Community Treatment and Service Center grant (Category III) grant from the Substance Abuse and Mental Health Services Administration (SAMHSA) to form the Philadelphia Alliance for Child Trauma Services (PACTS; Beidas, Adams et al., 2016). PACTS is a community academic partnership, including policymakers, leadership from community mental health agencies, and university-based researchers, whose mission is to increase the number of children who receive evidence-based trauma treatments in Philadelphia by (a) developing an integrated system of care for child trauma providers, (b) building capacity for trauma screening and assessment, (c) building partnerships between PACTS providers and other child serving systems (e.g., schools, child welfare, juvenile justice), and (d) increasing the delivery of evidence-based practices for trauma, with a particular focus on TF-CBT. Since 2012, PACTS has supported the training of eight cohorts of community behavioral health clinicians in TF-CBT across outpatient community mental health and residential treatment agencies. Training included 2 days of didactics and ongoing consultation provided via bi-weekly consultation calls for 8 months with a TF-CBT certified master trainer (See Beidas, Adams et al. (2016) for more details on PACTS and TF-CBT training).

Study Participants and Procedure

Study procedures were approved by the City of Philadelphia Institutional Review Board. All youth receiving TF-CBT through PACTS between 2013 and 2016 were eligible. Exclusion criteria included inability to assent or lack of availability of a legal guardian to provide consent (assents and consents were available in English and Spanish). Although we were not able to ascertain the exact number of youth eligible for the evaluation, we calculated an estimate by summing the number of youth who received TF-CBT in PACTS agencies per the monthly reports that supervisors from each agency sent to the PACTS coordinator employed by DBHIDS. Based on these estimates, approximately 440 youth received TF-CBT and were eligible for the evaluation between 2013 and 2016.

Each week, the evaluation coordinator sent an email to all PACTS clinicians and supervisors reminding clinicians to introduce the evaluation study to all youth receiving TF-CBT, who were within the first four sessions of TF-CBT. When a clinician identified a potential participant, the evaluation coordinator worked with the clinician to acquire written permission (i.e., “consent-to-contact”) from
the youth client and the client’s guardian for the evaluation team to contact them about further participation in the study. After the clinician faxed the “consent-to-contact” form to the evaluation coordinator, the coordinator then contacted the family to schedule a baseline visit. The client and guardian provided written assent and consent before completing the pre-treatment evaluation, which was conducted by trained interviewers in the agency where the youth was receiving services. For all assessment measures in our battery, the measures were interviewer administered with the youth (≥11 years old) or the caregiver (youth < 11 years old).

Participants included 114 clients receiving TF-CBT from clinicians employed by 15 PACTS agencies (range of clients per agency = 1–24; M clients per agency = 11; SD = 7; median = 7) from which either the youth or the caregiver completed the baseline visit assessment battery (i.e., 114 families consented of the approximately 440 families receiving TF-CBT who were eligible). Assessments occurred at baseline and every 6 months thereafter until treatment termination, when there was a final assessment (i.e., no after treatment follow-up assessments were conducted). Of the 114 youth who completed the baseline assessment battery, 63 (55% of baseline) completed the termination assessment. An additional nine youth who did not complete the termination assessment completed the 6-month follow-up assessment. In the current evaluation, our follow-up assessment is the last completed non-baseline assessment (i.e., we use the termination follow-up data, or the 6-month follow-up data if available and when the termination follow-up data were not available), consistent with previous research (Lang, Randall, Delaney, & Vanderploeg, 2017), which provides a total of 72 youth at follow-up. Thus, in the current evaluation, we had a baseline and follow-up assessment for 63% of youth participants, of those, 12.5% are the 6-month evaluation data.

Measures

**Child PTSD Symptom Scale**

The Child PTSD Symptom Scale (CPSS; Foa, Johnson, Feeny, & Treadwell, 2001) is a 24-item measure that assesses DSM-IV PTSD symptoms, as well as functional impairment due to PTSD symptoms. The PTSD Symptom Severity subscale includes the sum of 17 items rated on a 4-point Likert scale (0 = “Not at all or only one time” to 3 = “5 or more times a week/almost always”). The CPSS Functional Impairment subscale includes the sum of seven items rated on a dichotomous scale (yes/no?). Higher scores on the CPSS Symptom Severity and Functional Impairment subscales are indicative of greater PTSD symptom severity and functional impairment, respectively. Per Foa et al. (2001), a clinical cutoff of 11 on the PTSD Symptom Severity subscale has adequate sensitivity and specificity to discriminate a PTSD diagnosis status obtained via a structured clinical interview. The CPSS displays good internal consistency (Symptom Severity: \( \alpha = .89–.90 \); Functional Impairment: \( \alpha = .89 \)) and reliability (Symptom Severity: \( r = .74–.84 \); Functional Impairment: \( r = .70 \)) (Foa et al., 2001; Nixon et al., 2013), and in the current evaluation (youth CPSS Severity \( \alpha = .87 \), CPSS functioning \( \alpha = .77 \); parent CPSS Severity \( \alpha = .86 \), CPSS functioning \( \alpha = .74 \)).

**Ohio Mental Health Consumer Outcomes System—Ohio Youth Problem, Functioning, and Satisfaction Scales**

The Ohio Mental Health Consumer Outcomes System—Ohio Youth Problem, Functioning, and Satisfaction Scales (Ohio Scales; Ogles, Melendez, Davis, & Lunnen, 2001) include 48-items that evaluate problem severity, functioning, hopefulness, and satisfaction with mental health services from the perspective of youth and/or their caregiver. Primary outcomes were the Problem Severity and Functioning subscales to capture the youths’ more general mental health difficulties and functional impairment. The Problem Severity subscale is comprised of the sum of 20 items rated on a six-point rating scale (0 = “Not at All” to 5 = “All of the Time”) focusing on commonly reported problems by youth who receive behavioral health services, with higher scores indicative of more problems. The Functioning subscale is comprised of the sum of 20 items rated on a five-point rating scale (0 = “Extreme Troubles” to 4 = “Doing very well”) designed to rate the youth’s level of functioning in a variety of areas of daily activity (e.g., interpersonal relationships, recreation, self-direction, and motivation), with higher scores indicative of better functioning. The Ohio Scales display good internal consistency and reliability (youth \( \alpha = .90–.95 \), \( r = .72 \); parent \( \alpha = .95–.97 \) \( r = .88 \)) (Ogles et al., 2001), and internal consistency in the current evaluation (youth \( \alpha = .88 \); parent \( \alpha = .89 \)).

**Analytic Method**

We conducted three sets of analyses. First, we used descriptive analyses to capture participant characteristics. Second, we estimated the effectiveness of TF-CBT from baseline to follow-up. Given the nested structure of our data (i.e., youth nested within agencies), we evaluated whether we needed to account for nesting in our analyses using a multilevel modeling analytic approach. Intra-class correlation coefficients (ICC) suggested that agency characteristics explained very little of the variance in outcomes (ICCs < 0.02); thus, we analyzed the effectiveness of TF-CBT using linear regressions in Mplus with full
Descriptive sample statistics are presented in the PACTS column of Table 1. On average, youth were 12 years old (range 5–19) at the time of their first evaluation, and 57% were female. Youth participating in the evaluation largely identified as African/African American/Black (45%) followed by biracial (27%), Latinx (24%) and Caucasian/European/White (4%). We did not formally assess family income; however, all youth served by PACTS agencies are Medicaid recipients or uninsured. Nine percent of youth did not live with their primary caregiver and were homeless or lived in a residential treatment facility. There were no differences by demographic characteristics at baseline for those who participated in the follow-up evaluation in comparison to those missing at follow-up.

Effectiveness Outcomes

Using linear regression modeling to estimate the effectiveness of TF-CBT provided in PACTS agencies on youth symptom changes from start of treatment to follow-up, we found small, statistically significant decreases in PTSD Symptom Severity, as well as PTSD Functional Impairment (See Table 2) as measured by the CPSS. Although we found a small, statistically significant decrease in Problem Severity from baseline to follow-up, there was no change in General Functioning as measured by the Ohio Scales. To determine if these statistically significant changes were clinically significant, we calculated Reliable Change Indices (Jacobson & Truax, 1991) for the subset of youth who completed both the baseline and follow-up assessments. Thirty-three percent of youth had reliable, clinically significant improvements in their PTSD Symptom Severity (17 youth out of 51 with complete data), 29% of youth had reliable, clinically significant improvements in their PTSD Functional Impairment (14 youth out of 49 with complete data), and 48% of youth had reliable, clinically significant improvements in their Problem Severity (24 youth out of 50 youth with complete data).

Benchmarking

Comparison of this Sample with Samples from Efficacy and Effectiveness Studies

Descriptive sample statistics from the PACTS evaluation participants, the efficacy study (Cohen, Deblinger et al., 2004), and the effectiveness study (Jensen et al., 2014) are presented in Table 1. In the PACTS evaluation, we recruited youth with a wider range of ages in comparison to the youth in the other two studies, including significantly older youth than in the Cohen, Deblinger et al. (2004) study and significantly younger youth than in the Jensen et al. (2014) study. We recruited significantly more males in comparison to both the Cohen, Deblinger et al. (2004) and Jensen et al. (2014) studies where the vast majority of participants were female. The PACTS youth were also more racially and ethnically diverse in comparison to both the Cohen, Deblinger et al. (2004) and Jensen et al. (2014) studies. We are unable to directly compare family incomes across studies, but through visual inspection of demographics, youth in the Jensen et al. (2014) study were from more affluent families in comparison to youth in the

Results

Preliminary Analyses

Descriptive sample statistics from the PACTS evaluation participants, the efficacy study (Cohen, Deblinger et al., 2004), and the effectiveness study (Jensen et al., 2014) are presented in Table 1. In the PACTS evaluation, we recruited youth with a wider range of ages in comparison to the youth in the other two studies, including significantly older youth than in the Cohen, Deblinger et al. (2004) study and significantly younger youth than in the Jensen et al. (2014) study. We recruited significantly more males in comparison to both the Cohen, Deblinger et al. (2004) and Jensen et al. (2014) studies where the vast majority of participants were female. The PACTS youth were also more racially and ethnically diverse in comparison to both the Cohen, Deblinger et al. (2004) and Jensen et al. (2014) studies. We are unable to directly compare family incomes across studies, but through visual inspection of demographics, youth in the Jensen et al. (2014) study were from more affluent families in comparison to youth in the
Cohen, Deblinger et al. (2004), and our study, with 87% of families in the Jensen sample with a family income of over $87,000 per year in American dollars.

Comparison of PACTS Study Characteristics with Study Characteristics from Efficacy and Effectiveness Studies

Descriptive study characteristics from the PACTS effectiveness study, efficacy study participants (Cohen, Deblinger et al., 2004), and the Norway effectiveness study participants (Jensen et al., Descriptive study characteristics from 2014) are presented in Table 3. Youth treated in the PACTS and Jensen et al. (2014) studies were treated in similar non-problem specific community clinics that served youth with heterogeneous types of trauma, whereas in the Cohen, Deblinger et al. (2004) study, all youth experienced sexual abuse and were being treated in a trauma-specific clinic. The length of clinicians’ initial

![Table 1 Descriptive characteristics and comparisons of participants in studies](image)

![Table 2 PACTS clinical outcomes](image)

Results were analyzed in Mplus with demographic auxiliary variables and Full Information Likelihood Estimation to account for attrition. RCI, Reliable Change Index; SD, Standard Deviation; SE, Standard Error.
training in TF-CBT varied across the different studies, but the format was similar. PACTS clinicians were provided with the least number of training days (2 days) and the Jensen et al. (2014) study clinicians the most (4–6 days). Clinicians in the PACTS evaluation and Jensen et al. (2014) study also completed a publicly available web-based learning course for TF-CBT (http://www.musc.edu/tfcbt). Across the three studies, all study clinicians received weekly supervision, and those in the Cohen, Deblinger et al. (2004) study and the PACTS evaluation also received bi-weekly consultation phone calls from TF-CBT trainers. In the PACTS evaluation and Cohen, Deblinger et al. (2004) study, the majority of clinicians were Masters-Level clinicians (J.A., Cohen, personal communication), but in the Jensen et al. (2014) study the majority of clinicians were Doctoral-Level psychologists.

Cohen, Deblinger et al. (2004) included the strictest exclusionary criteria. Youth in the Cohen, Deblinger et al. (2004) and Jensen et al. (2014) studies were evaluated for follow-up at specific sessions (i.e., 12th for Cohen, 15th for Jensen), while those in the PACTS evaluation were evaluated every 6 months and their last available data point was used for follow-up. Youth in the PACTS evaluation had the highest level of attrition (32%), followed by the Jensen et al. (2014) study (22%). There were no baseline demographic or symptom differences between youth

Table 3 Study characteristics

<table>
<thead>
<tr>
<th>Study characteristics</th>
<th>PACTS</th>
<th>Cohen, Deblinger et al. (2004)</th>
<th>Jensen et al. (2014)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of trauma treated</td>
<td>Heterogeneous</td>
<td>Sexual abusea</td>
<td>Heterogeneous</td>
</tr>
<tr>
<td>Study context</td>
<td>Community clinics</td>
<td>Trauma clinics</td>
<td>Community clinics</td>
</tr>
<tr>
<td>Training in TF-CBT</td>
<td>Two days of didactic training; web-based course; 8 months of bi-weekly consultation; weekly supervision</td>
<td>Three days of training in both TF-CBT and Child-centered Therapy; Weekly supervision; twice monthly cross-site phone supervision</td>
<td>Four to six days of initial training in TF-CBT; web-based course; weekly supervision that reduced to bi-weekly as therapist became more familiar with treatment</td>
</tr>
<tr>
<td>Therapists</td>
<td>Masters level clinicians</td>
<td>75%–80% social workers; 20%–25% psychologist</td>
<td>81% psychologists; 8% psychiatrists; 8% educational therapists; 4% social workers</td>
</tr>
<tr>
<td>Comparison group</td>
<td>–</td>
<td>Child-centered therapy</td>
<td>Treatment as usual</td>
</tr>
<tr>
<td>TF-CBT provided</td>
<td>–</td>
<td>12 weekly individual sessions to parent-child dyad; 90 minutes total, 45 to each member of dyad; three joint parent-child session</td>
<td>12–15 weekly sessions to parent-child; parallel and co-joint sessions</td>
</tr>
<tr>
<td>Inclusionary criteria</td>
<td>1. Receiving therapy from a therapist trained in TF-CBT</td>
<td>1. 5 + criteria for PTSD, including one symptom in each of the three PTSD clusters</td>
<td>1. Experienced at least one traumatizing event</td>
</tr>
<tr>
<td></td>
<td>2. Within the first four sessions of treatment</td>
<td>2. Caretaker willing to participate in the parental treatment component of the study</td>
<td>2. CPSS score of 15+ and at least one symptom in each of the three PTSD clusters</td>
</tr>
<tr>
<td>Exclusionary Criteria</td>
<td>1. No guardian to consent</td>
<td>1. Acute psychosis</td>
<td>1. Acute psychosis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Parent or child with a functionally impairing substance abuse disorder</td>
<td>2. Suicidal behavior</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Documented developmental disorder</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. Children on psychotropic medication had to have been stable on medication for at least 2 months</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6. Could not be receiving psychotherapy for sexual abuse outside of the study</td>
<td></td>
</tr>
<tr>
<td>Follow-up assessment</td>
<td>Last assessment available</td>
<td>After the 12th session</td>
<td>After the 15th session</td>
</tr>
<tr>
<td>Attrition (% participation at follow up)</td>
<td>68</td>
<td>88</td>
<td>78%</td>
</tr>
<tr>
<td>Demographic differences by attrition</td>
<td>None</td>
<td>None</td>
<td>1. Age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. Number of traumas</td>
</tr>
</tbody>
</table>

aWhile they were in treatment for their recent sexual abuse, children reported experiencing 2.66 (SD = 1.61) additional traumas prior to the intake. Endashes indicate that the information was not available in the manuscript.
who were present versus missing for follow-up in the PACTS evaluation and Cohen, Deblinger et al. (2004) study; however, in the Jensen et al. (2014) study, older youth and those exposed to a greater number of traumas were more likely to drop out of treatment.

Comparison of PACTS Clinical Outcomes with Clinical Outcomes from Controlled Efficacy and Effectiveness Studies

See Table 4 for the benchmarked clinical outcomes across the PACTS, Cohen, Deblinger et al. (2004), and Jensen et al. (2014) studies. The Kiddie-Schedule for Affective Disorders and Schizophrenia for School-Aged Children-Present and Lifetime Version Re-experiencing, Avoidance, and Hypervigilance subscales (KSADS; Kaufman et al., 1997) was used to measure PTSD symptom severity in the Cohen, Deblinger et al. (2004) study. In both the PACTS evaluation and the Jensen et al. (2014) effectiveness study, the CPSS was used to measure changes in PTSD symptoms. As such, we can statistically compare the raw scores for the CPSS Symptom Severity score at baseline and follow-up; however, we cannot compare the CPSS Functional Impairment score because different versions of this subscale were used. One of the inclusionary criteria for the Jensen et al. (2014) study was a CPSS Symptom Severity score above 11; thus, 100% of youth in the Jensen et al. (2014) study had a CPSS Symptom Severity score above 11, which was significantly higher than the 87% of youth with a CPSS Symptom Severity score of above 11 at baseline in the PACTS evaluation. As such, youth in the Jensen et al. (2014) study had significantly higher average CPSS Symptom Severity scores at baseline in comparison to the youth in the PACTS evaluation. At follow-up, significantly more youth from the Jensen et al. (2014) study had CPSS Symptom Severity scores below 11 (68%) than PACTS youth (46%). Youth in the PACTS evaluation had significantly higher average CPSS Symptom Severity scores at final assessment in comparison to youth in the Jensen et al. (2014) study.

To compare the effects of TF-CBT across the three studies, we converted the means and standard deviations that were available in the manuscripts for the baseline and follow-up assessments into Cohen’s ds with associated 95% confidence intervals. As mentioned above, PTSD symptom severity (but not the functional impairment score) was measured using the same version of the CPSS for the PACTS evaluation and the Jensen et al. (2014) study, so the effect sizes can be directly compared. The effect sizes for PTSD symptom and functional improvement among the youth in PACTS evaluation were small ($d = 0.34$, 95% CI: 0.08, 0.60) and significantly smaller “due to non-overlapping confidence intervals,” than the large effect sizes for youth in the Jensen et al. (2014) effectiveness study ($d = 1.69$, 95% CI: 1.28, 2.07). The effect size for PTSD symptom improvement among the youth in the Cohen, Deblinger et al. (2004), as measured by the K-SADS, was large. General Mental Health was measured using the Ohio Problem Severity score in the PACTS evaluation, the Child Behavioral Checklist (Achenbach & Rescorla, 2001) in the Cohen, Deblinger et al. (2004), and the Strengths and Difficulty Questionnaire (Goodman, 1997) in the Jensen et al. (2014) study. The effect size for general mental health among youth in the PACTS evaluation was small, medium in the Cohen, Deblinger et al. (2004) study, and large in the Jensen et al. (2014) study.

Discussion

The goal of the current study was to preliminarily evaluate the effectiveness of TF-CBT delivered by community clinicians in an urban setting that primarily provided services to youth living in poverty by benchmarking clinical outcomes against published efficacy and effectiveness studies. We found that youth who received TF-CBT in community behavioral health agencies in the city of Philadelphia had modest improvements in trauma symptoms, functional impairment due to their trauma symptoms, and their overall general mental health. However, the magnitude of the improvements was significantly smaller than those found in the Cohen, Deblinger et al. (2004) efficacy RCT and the Jensen et al. (2014) effectiveness RCT, and significantly fewer youth in this sample demonstrated non-clinical PTSD symptoms at the end of treatment. These findings are in contrast to recent statewide effectiveness trials of TF-CBT provided in community settings in Connecticut (Donisch, 2018) and Delaware (Webb et al., 2014) which found that youth who received TF-CBT had large improvements in their PTSD symptoms and small to medium improvements in their comorbid symptoms (e.g., internalizing, externalizing, general mental health), similar to the efficacy RCT (Cohen, Deblinger et al., 2004) and effectiveness RCT (Jensen et al., 2014) study findings.

Limitations

The current findings must be interpreted within the context of several limitations. First, there were significant missing data at follow-up. Only 63 of 114 youth completed the termination evaluation. This is primarily due to significant treatment dropout. To enhance our follow-up sample size, we included the six-month follow-up data when available, and when termination data were missing.
and used FIML estimation (i.e., a statistical method that accounts for missing data). However, it is important to interpret the findings with the understanding that 12.5% of the follow-up sample had not yet completed the entire protocol of TF-CBT, which could attenuate the effect sizes of our findings. We conducted a sensitivity analysis among the 63 youth who completed the termination evaluation and found similar small effect sizes as presented above. Second, a limitation of this study was our inability to interpret the findings. There are numerous challenges to obtaining fidelity assessments in community settings, and it is critical to identify accurate and cost-effective way to assess fidelity when evidence-based practices are being naturalistically implemented in the community (Beidas, Maclean et al., 2016). However, as we were unable to assess fidelity in this study, it is unclear whether clinicians were implementing TF-CBT as intended, or if the small effect sizes were due to lack of fidelity. Third, a potential limitation of our study is the generalizability of the results to the larger population of urban youth. Only one in four of eligible youth consented to be in the study. Reasons for non-consent include the clinicians not introducing the study, youth feeling uncomfortable with participation, and unavailability of a caregiver to provider consent. Fourth, a limitation of the study is the variability of reporters for our clinical outcomes. Because clinicians served a wide age range of children (5–19), a portion of our clinical outcome measures were completed by parents, rather than youth. We collapsed across reporter, again, to enhance our sample size, and because we were not powered to evaluate outcomes by reporter. Finally, given the complex traumas experience by youth and conversations with our stakeholders, we prioritized reducing participant fatigue. Thus, there were several measures we were unable to include in our assessment battery (e.g., number/type of traumas experienced, adverse childhood events, and experiences with discrimination) that could have further informed plans for adaptation. Taking into account these limitations, there are several potential explanations for why youth in this study had smaller clinical improvements to previous TF-CBT. These explanations are derived from collective perspectives from the PACTS community academic partnership. These preliminary findings suggest that there may be room for improvement.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CPS differential (95% CI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% with score above 11</td>
<td>Baseline: 85%,</td>
<td>–</td>
<td>Baseline: 100%,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n = 85/100</td>
<td></td>
<td>n = 79/79</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Follow Up: 68%,</td>
<td>–</td>
<td>Follow Up: 46%,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n = 40/59</td>
<td></td>
<td>n = 27/59</td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>Baseline: 23.82(11.66), n = 114</td>
<td>–</td>
<td>Baseline: 26.82 (8.05), n = 79</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Follow Up: 19.70(12.38), n = 114</td>
<td>–</td>
<td>Follow Up: 11.34 (10.52), n = 59</td>
<td>–4.42a –12.09, –4.63</td>
</tr>
<tr>
<td>Cohen's d (95% CI)</td>
<td>0.34 (0.08, 0.60)</td>
<td>–</td>
<td>1.69 (1.28, 2.07)</td>
<td></td>
</tr>
<tr>
<td>CPS differential (95% CI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>Baseline: 2.84 (2.11), n = 114</td>
<td>–</td>
<td>Baseline: 8.03 (1.84), n = 79</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Follow Up: 2.08 (1.89), n = 114</td>
<td>–</td>
<td>Follow Up: 10.33 (1.99), n = 58</td>
<td>–</td>
</tr>
<tr>
<td>Cohen's d (95% CI)</td>
<td>0.38 (0.12, 0.64),</td>
<td>–</td>
<td>–1.21 (–1.57, –0.83)</td>
<td>–</td>
</tr>
</tbody>
</table>
| Kiddie-schedule for affective disorders and schizophrenia for school-aged children-present and lifetime version
  Reexperiencing Cohen's d (95% CI)                    | –              | 1.82 (1.49, 2.14)              | –                    | –                      |
  Avoidance Cohen's d (95% CI)                         | –              | 1.73 (1.40, 2.04)              | –                    | –                      |
  Hypervigilance Cohen's d (95% CI)                    | –              | 1.6 (1.27, 1.91)               | –                    | –                      |
| General mental health                                |                |                                |                      |                                  |
  Ohio problem severity Cohen's d (95% CI)             | 0.29 (0.03, 0.55)| –                              | –                    | –                      |
  CBCL Total Cohen's d (95% CI)                        | –              | 0.67 (0.38, 0.95)              | –                    | –                      |
  SDQ Cohen's d (95% CI)                               | –              | 1.24 (0.86, 1.61)              | –                    | –                      |

Cohen’s d calculated from baseline to follow up. Endashes indicate that the information was not reported in manuscript or cannot be calculated.

95% CI, 95% Confidence Interval; CBCL, Child Behavior Checklist; CPSS, Child PTSD Symptom Scale; SD, Standard Deviation; SDQ, Strengths and Difficulties Questionnaire.

*p < .05.

Table 4 Benchmark outcomes
with regards to providing trauma treatment for youth in urban setting who are living in poverty (Ruchkin, Henrich, Jones, Vermeiren, & Schwab-Stone, 2007).

Interpretation of Results

By benchmarking the PACTS evaluation results to the Cohen, Deblinger et al. (2004) efficacy RCT and the Jensen et al. (2014) effectiveness RCT, we are able to compare our treatment outcomes to precise treatment effects, as well as treatment effects found in the community aligned with traditional efficacy and effectiveness study research methods. There were notable differences in the study characteristics between the three studies, with the most dramatic including participant inclusionary and exclusionary criteria. In the efficacy RCT (Cohen, Deblinger et al., 2004), great care was taken to ensure that participants had a diagnosis of PTSD and did not have other comorbid disorders (e.g., substance abuse disorders, developmental disorders, psychosis), whereas in the effectiveness RCT study (Jensen et al., 2014), these requirements were less strict (i.e., CPSS score of 15+ and at least one symptom from each of the three PTSD clusters). In our study, there were no exclusion criteria related to comorbidities. While we do not know the full diagnostic profile of youth receiving treatment in this sample, in Beidas, Adams et al. (2016), per Medicaid claims data, youth treated with TF-CBT in PACTS agencies had heterogeneous diagnoses, including disruptive behavior disorders and mood disorders in addition to PTSD. It is possible that the small effect sizes found in this sample are due to more complex diagnostic profiles in presenting youth, which could have attenuated the effectiveness of TF-CBT.

There may have been differences in the skill with which TF-CBT was provided across the efficacy and effectiveness studies. PACTS clinician training in TF-CBT included two days of didactic training and eight months of bi-weekly consultation calls with a TF-CBT certified trainer (Beidas, Adams et al., 2016). Further, each agency had a designated internal TF-CBT supervisor for weekly supervision of cases, and there were annual booster trainings to keep clinicians engaged in TF-CBT implementation. While clinicians across the three studies received similar levels of initial training, clinicians in this study were allowed to start recruiting participants for the evaluation before completing the training requirements for certification (i.e., they may not have completed their certification cases), whereas clinicians in the efficacy RCT (Cohen, Deblinger et al., 2004) were required to be certified TF-CBT clinicians prior to recruiting clients for their studies. Further, the clinicians in the efficacy RCT (Cohen, Deblinger et al., 2004) were closely supervised by the developers of the intervention. Thus, differences in effect sizes across studies could be due to the skill with which TF-CBT was provided as there were differing requirements for certification with TF-CBT at the time of evaluation.

Meta-analyses find that implementing evidence-based practices with fidelity is related to larger treatment outcomes (Durlak & DuPre, 2008). However, while most clinicians rate their likelihood of using all TF-CBT components as high, only 66% of community clinicians regularly implement all components of TF-CBT (Allen & Johnson, 2012). As mentioned above, we were unable to directly assess fidelity to TF-CBT components in this study, but during informal conversations, clinicians noted difficulties with caregiver engagement and completing the joint parent–child sessions. Joint parent–child sessions are a critical component of the intervention (Cohen, Deblinger et al., 2004). Given the high comorbidity of disruptive behavior disorders within PACTS youth broadly (Beidas, Adams et al., 2016), the parent education components were likely even more crucial. These difficulties are consistent with the larger literature regarding caregiver engagement among low-income and at-risk youth. A previous study found that incorporating parental engagement strategies into TF-CBT increased the likelihood that foster parent–child dyads finished the first four sessions of TF-CBT and completed the entire treatment protocol in comparison to traditional TF-CBT (Dorsey, Kerns, Trupin, Conover, & Berliner, 2012). Thus, it may be critical for TF-CBT clinicians who work with low-income urban youth to be trained in how to implement effective parent engagement strategies in order to optimize TF-CBT through parent involvement (Staudt, 2007).

TF-CBT was originally developed for youth who experienced sexual trauma (Cohen & Mannarino, 1996, 1997, 1998; Deblinger, Lippmann, & Steer, 1996), but was later found to improve symptoms related to other traumas such as catastrophes (Hoagwood et al., 2007) and traumatic grief (Cohen, Mannarino, & Staron, 2006; Cohen, Mannarino et al., 2004). An assumption of the original treatment protocol is that the trauma is in the past and the child can process, as well as move forward, from the trauma. However, many youth, particularly youth living in low resourced urban settings, experience ongoing traumas such as community violence. Youth in Philadelphia often experience multiple adversities before age five, and youth in our study were living in low-income neighborhoods in Philadelphia with high risk for continued exposure to community violence and other adversities (e.g., Last et al., in preparation; Pew, 2017). Cohen, Mannarino, and Murray (2011) provided recommendations for adaptations for this population, including addressing safety early in
treatment rather than at the end, and adapting the trauma narrative processing to differentiate between real danger and generalized trauma reminders. It is critical that TF-CBT providers who treat youth in urban settings at risk for experiencing ongoing traumas are trained in these adaptations. Further, it is important for clinicians to have the tools and training to make decisions regarding which evidence-based practice or adapted evidence-based practice they will use with each client. However, it should be emphasized that, while much has been written about how to make these adaptations, these adaptations have not been rigorously evaluated to determine if they improve the effectiveness of the TF-CBT for chronically traumatized youth.

Moreover, PACTS youth were significantly more diverse than the youth represented in the efficacy RCT (Cohen, Deblinger et al., 2004) and effectiveness RCT (Jensen et al., 2014). The majority of youth served were from marginalized groups. The experience of discrimination is consistently linked to poorer mental health, particularly the experience of trauma-like symptoms (Pascoe & Smart Richman, 2009). It is critical that TF-CBT is delivered in a manner that is sensitive to economic disparities, racism, and historical traumas that the youth may experience on a daily basis. Finally, all youth served in this study were low-income and on Medicaid. This may highlight the first need to have their basic needs met. This stresses the need for case management, resource allocation, care coordination for chronically traumatized youth (Dorsey et al., 2012).

Implications and Future Directions

Given the high rates of traumatic events that urban, low-income, and ethnic/racial minority children and youth in the United States face, it is critical for high-quality treatments for PTSD to be available in community settings. While TF-CBT is an intervention with a large literature supporting its efficacy for improving PTSD symptoms, as well as more general mental health outcomes, there is limited literature supporting its effectiveness among urban youth experiencing heterogeneous, chronic, and comorbid traumas in community clinics. While this preliminary study found that youth treated with TF-CBT in the community improved, the small effects are worth additional study, and consideration, particularly given national efforts to widely implement TF-CBT. The following are important future research directions to ensure that we are providing evidence-based and effective treatments for PTSD to youth in our communities. Future research should:

1. Evaluate the effectiveness of TF-CBT, and other psychotherapies for trauma symptoms, implemented within community mental health clinics under naturalistic conditions.
2. Evaluate the effectiveness of policies that aim to improve the contexts within which urban youth live. Effective trauma treatments are necessary but not sufficient for urban youth. Systematic policies are needed to invest in the economic and social growth of urban communities.
3. Develop simple methods for assessing fidelity to interventions to inform effectiveness research efforts (Beidas, Maclean et al., 2016).
4. Develop assessment protocols to determine which intervention (or combination of interventions if TF-CBT plus parent engagement strategies or care coordination is needed) is most likely to be effective for youth, and better implement the most effective treatments, in the most helpful sequence, for differing populations of youth.
5. Evaluate the adaptations that clinicians must make to interventions to fit their organization and community context, as well as proposed adaptations to address populations which face more chronic or severe traumas (e.g., Cohen et al., 2011) to ensure that youth receive an evidence-based intervention, and, at a minimum, ensure that there are no iatrogenic effects.

As we continue to study the implementation of evidence-based interventions within community organizations, it is critical that we simultaneously evaluate the effectiveness of these interventions. Implementation efforts should inform treatment development and revision, just as treatment development and adaptation should inform future implementation efforts.

ACKNOWLEDGMENTS We are especially grateful for the support and partnership that the Department of Behavioral Health and Intellectual disAbility Services provided for this project and for the Evidence Based Practice and Innovation (EPIC) group and Ronnie Rubin, PhD. We thank Dr. Judy Cohen for her input on this manuscript. We are also very grateful to the individuals who have been a part of the Philadelphia Alliance for Child Trauma Services (PACTS), including the clinicians, supervisors, administrators, partners, and families. Brittany Rudd was supported by a National Institute of Mental Health Training Fellowship (T32 MH109433). Funding for this project was supported by grants from the Substance Abuse and Mental Health Administration (SM61087 Jackson and Berkowitz), and the National Institute of Mental Health (K23 MH099179, Beidas).

References


