

Outcomes of Wraparound Care Coordination for Youth with Complex Behavioral Health Needs

Spencer Hensley

A thesis

Submitted in partial fulfillment of the
requirements for the degree of

Master of Science

University of Washington

2019

Committee:

Ann Vander Stoep

Eric J Bruns

Program Authorized to Offer Degree:
Epidemiology

©Copyright 2019

Spencer Hensley

University of Washington

Abstract

Outcomes of Wraparound Care Coordination for Youth with Complex Behavioral Health Needs

Spencer Hensley

Chair of Supervisory Committee:

Ann Vander Stoep

Epidemiology

Background

Approximately 10% of youth in the US are estimated to experience “serious emotional disorder,” or SED. Wraparound is a commonly implemented care coordination process for youth with SED and complex needs, but effectiveness research is sparse, and results have been mixed. RCTs have found favorable effects of Wraparound on child welfare placements and service utilization in their samples, but only two RCTs have measured youth emotional and behavioral functioning outcomes directly, and they have found null results. Thus, the need for more evidence about the effectiveness of Wraparound on youth functioning remains. This study uses a propensity-score-matched comparison group to measure the effectiveness of Wraparound on youth functioning across multiple areas of need.

Methods

Data from this study come from a large 501(c)(3) non-profit behavioral health provider agency with multiple locations in California. All youth who received services through this agency between October 1, 2015 and December 17, 2018 and who completed assessments of emotional and behavioral functioning at intake and six-months after enrollment were eligible for inclusion in the study. A total of 129 Wraparound-enrolled youth were included. These youth were propensity-score-matched on 13 relevant demographic and clinical acuity variables to youth who received services other than Wraparound (N = 1,154 youth). After matching, 122 Wraparound-enrolled and 122 comparison youth were included for analyses.

Youth functioning was measured using the Child and Adolescent Needs and Strengths (CANS), an assessment that rates youth functioning with respect to the “actionability” of discrete needs. We measured change in three ways: the net change in number of needs considered “actionable,” the number of needs resolved after six months, and the number of new needs after six months. Matching successfully balanced youth with respect to baseline functioning. Linear regressions and negative binomial regressions were used to compare change scores between the Wraparound-enrolled and comparison youth.

Results

Wraparound youth experienced a greater net reduction in actionable needs after six months (Mean Difference: -0.36, 95%CI: -2.03, +1.32), though confidence intervals were wide such that this difference was non-significant. Similarly, Wraparound youth experienced both a greater reduction in actionable needs (IRR: 1.32, 95%CI: 0.98, 1.79) and had new needs discovered more often than non-Wraparound youth (IRR: 1.23, 95%CI: 0.92, 1.65), though neither of these outcomes differed significantly between the two groups.

Conclusions

Our results contribute to the growing effectiveness research base on the behavioral, emotional, and functional impacts of Wraparound. Although the estimated effects of Wraparound in this study only approached significance, between-group differences were clinically meaningful as per conventions for interpreting results from the CANS. The resolution and discovery of more needs after six months may be explained by the additional and intensive time spent with families by Wraparound provider staff.

INTRODUCTION

The National Research Council and Institute of Medicine estimate that between 14% and 20% of youth in America have a diagnosable mental health disorder¹. Some subset of these youth can be described as experiencing “serious emotional disorder” (SED; also called “serious emotional disturbance”), a term defined in the *Federal Register* as having a diagnosable mental, behavioral, or emotional disorder that results in “substantial” interference with family, school, or social functioning². A recent meta-analysis of twelve prevalence studies of SED estimates a pooled prevalence of SED with impairment in family, peer, educational, or community settings of 10.0% and SED with impairment across more than one of these domains of 6.4%.³ Youth with SED are typically the most impaired among youth with mental disorders. They often have complex problems, navigate multiple service systems (e.g., mental health, child welfare, juvenile justice, special education) and have multiple professional “helpers” in their lives. The approximately 10% of youth who are the highest utilizers of behavioral health services in Medicaid account for approximately 40% of all youth behavioral health service costs⁴.

Wraparound is a care coordination process for youth with serious emotional disorders and complex needs⁵. Wraparound is individualized, family-driven, and community-based. Youth and families are typically part of a Wraparound team that includes both professional supports, such as mental health clinicians and representatives from relevant child-serving agencies, and natural supports, such as friends and extended family. A professional care coordinator works together with a youth, her family, and the team to develop a comprehensive and individualized plan of services and strategies to meet the needs of the youth and family. Often, peer support from experienced parents or other youth are also available to the youth and family.

Effectiveness research about Wraparound is sparse⁶. To date, results of only 6 experimental studies of Wraparound have been published in peer reviewed journals in which the outcomes for Wraparound-enrolled youth are compared to a control group who did not receive Wraparound⁷⁻¹². Relying on administrative data, these studies have found favorable effects of Wraparound on child welfare placement^{7,9} and service utilization¹², and have found no effects on juvenile justice recidivism¹⁰. However, in the two experimental studies where youth emotional and behavioral functioning was measured more directly by standardized assessments, findings were null across all domains^{8,11}. Both of these studies had small samples, and in one of the studies implementation fidelity was markedly low¹¹. In the other, implementation fidelity and practice quality were not measured⁸.

Quasi-experimental research has shown more consistently favorable results, with Wraparound participants consistently demonstrating more positive outcomes, such as reduction in symptoms, improved functioning, and less restrictive residential placement, than comparison groups. Several studies have found improvements in behavioral health functioning as measured by self-report or standardized assessments^{6,13,14}. To our knowledge, two effectiveness studies have relied on propensity score matching (PSM) to create a comparison group—an approach used to reduce the inherent selection bias of observational studies^{15,16}. One study that relied on PSM calculated propensity scores that balanced the Wraparound and comparison group with respect to age, gender, psychiatric diagnosis, insurance type, and history of residential psychiatric admission. The authors found that Wraparound youth had comparatively lower claims expenses, utilized lower intensity services, and maintained less restrictive residential placements¹⁶. The Wraparound-enrolled youth also showed improvement in clinical functioning scores between 6% and 20% after twelve months, as measured by four different standardized

assessments. However, none of these standardized assessments was collected for the comparison group, so the additional impact of Wraparound on clinical functioning, relative to the comparison group, could not be estimated.

The other study that utilized PSM was a dissertation by Karpman. This study, conducted with Medicaid-eligible youth in Massachusetts, relied on a widely-used measure of youth behavioral health and functioning, the Child and Adolescent Needs and Strengths (CANS), to assess outcomes¹⁵. The author found that Wraparound-enrolled youth improved slightly but at a significantly slower rate than the control group. The author concluded that Wraparound-enrolled youth experienced poorer outcomes as assessed by the CANS. A recently published literature review highlighted this dissertation as the only published effectiveness study to find statistically worse outcomes for Wraparound youth⁶. However, the CANS is a complex measure of youth functioning, and there are many ways to calculate change with the tool. One methodological feature of Karpman's study is that they quantified change in functioning via total scores calculated across four broad domains. None of these four measures of change considered the independent nature of CANS items, which could be used to detect improvement in some facets of behavior or functioning and worsening in others.

There is a continued need for research about the behavioral and functional impacts of the Wraparound process on youth with serious and complex needs, including research on whether and how specific behavioral and functional domains may improve over time for Wraparound-enrolled youth. In addition, given the diversity of Wraparound initiatives nationally (e.g., with respect to the populations served, practice models employed, and types of implementation support), and relative paucity of rigorous outcomes studies, the research base will benefit from additional studies.

This thesis will follow a similar design to the Massachusetts dissertation study by Karpman. The current study employed a comparison group using propensity score matching, and the CANS was used as the outcome measure. Building upon the prior study, this study focuses on children and adolescents enrolled in mental health services in California and undertakes a more refined outcome analysis, estimating change over time in a way that (1) accounts for individual behavioral and functional domains assessed by specific CANS items, and (2) evaluates both improvement and worsening across these domains/items. By doing so, this thesis aims to contribute to the research base about Wraparound effectiveness and to determine if Wraparound-enrolled youth resolve and/or develop more behavioral and emotional needs, as measured by the CANS, than do similar youth who do not enroll in Wraparound.

METHODS

Study Design. We used a retrospective cohort study design to assess the effect of Wraparound on behavioral, emotional, and functional outcomes. We created a propensity-score-matched sample to act as a comparison group to youth who received Wraparound.

Study Setting. Data came from a large 501(c)(3) non-profit behavioral health provider agency with multiple locations in California. The agency provides a continuum of care for children and young adults ranging from 0-24 years old. Although precise numbers are not available, the agency estimates that approximately 95% of Wraparound-enrolled youth are Medicaid-enrolled.

Measures. The Child and Adolescent Needs and Strengths (CANS) assessment was used both for matching and as an outcome measure. The CANS is an assessment built around a

“communimetric” (as opposed to psychometric) approach to measurement that prioritizes the communication of actionable information for use in clinical decision-making¹⁷. Rather than symptoms of behavioral health problems, CANS items are individual “needs” that span the breadth of issues that may complicate children’s lives. Example items include “Self-Injurious Behavior,” “Sexual Aggression,” “Social Functioning,” “Anxiety,” and “Substance Use.” Each need can be scored between 0 and 3, where 0 indicates that there is “no evidence” of a need, a 1 indicates that “watchful waiting or prevention” is warranted, a 2 indicates a need that requires “action,” and a 3 indicates a need that requires “immediate or intensive action.” Together, 2s and 3s are considered “actionable” needs, or needs that are appropriate targets of behavioral health treatment. When dichotomizing items as actionable, CANS summative scores can be conceptualized as a count of “needs,” or a count of distinct areas of functioning where a youth would benefit from intervention.

In the specific version of the CANS used in the current provider agency, 79 items are organized into 7 domains: Behavioral and Emotional Needs (example items include “Psychosis” and “Depression”), Life Functioning (e.g., “Family Relationships,” “School Functioning”), Child Risk Behaviors (e.g., “Self-Injurious Behavior,” “Sexual Aggression”), Cultural Factors (e.g., “Language,” “Discrimination/Bias”) Caregiver Resources and Needs (e.g., “Involvement with Care,” “Knowledge”), Child Strengths (e.g., “Optimism” and “Community Life”), and Traumatic Adverse Childhood Experiences (e.g., “Physical Abuse” and “Medical Trauma”). For each of these items, small descriptions accompany each anchor. For example, the guidance for assigning a score of “3” to “Depression” reads, “Clear evidence of depression that is disabling for the child in multiple life domains.”

Some CANS items are collected in optional “modules” that provide additional detail about a need. These items are only completed by youth with an actionable need on specific triggering items from the domains. For example, youth for whom the item “Runaway” is an actionable need (scored 2 or 3) are also scored on runaway module that includes eight additional items: consistency of destination, frequency of running, involvement of illegal acts, involvement of others, likelihood of return on own, planning, realistic expectations, and safety of destination. Unlike the items in the domains, these module items do not represent discrete needs and are not completed about all youth. For these reasons, they are not included in the analyses for the current paper.

The CANS has demonstrated strong inter-rater reliability among both researchers and caseworkers¹⁸. However, it is a complex measure, and change in child outcomes can be calculated in several ways. The CANS is sometimes used as a way to determine eligibility for different types and/or intensities of services.^{19,20} The developers of the CANS work with licensees to customize the assessment, adding, removing, or modifying items. As a result, the CANS is not precisely the same assessment in any two contexts. No changes to the CANS were made during the period of time under study in this paper. See Appendix A for the complete assessment as used in this study.

We used baseline and six-month CANS assessments. A baseline CANS was defined as any CANS administered within 30 days before or after enrollment. A six-month CANS was defined as any CANS administered between 150 and 210 days after enrollment. Both baseline and six-month CANS were completed by the Wraparound care coordinator in collaboration with the youth and family.

Study Participants. Participants in the Wraparound group included all youth ages 5-18 who received Wraparound care coordination at the behavioral health provider agency from October 1, 2015 through December 17, 2018. Comparison group participants were identified using propensity score matching from a pool of youth served by the same behavioral health provider agency during the same years who received another type of service, i.e., foster-family based services, mobile crisis, kinship care, outpatient, school-based or residential mental health services. All youth who completed a baseline and six-month follow up assessment within the study time frame were eligible as study participants. A total of 461 Wraparound-enrolled youth completed a Baseline CANS assessment. Of these, 129 (28.0%) youth also completed their six-month CANS assessment and were eligible for inclusion in this study. The pool of potential comparison youth with complete baseline and six-month assessments was 1,154.

Compared to the full cohort of N=461 Wraparound-enrolled youth with baseline CANS data, Wraparound-enrolled youth who completed both a baseline and six-month assessment had fewer needs at baseline in two domains: Behavioral and Emotional Needs and Life Functioning. They had more needs at baseline in the Child Risk Behaviors domain. With respect to the other four CANS domains, as well as age, sex, and race, the Wraparound youth who completed a baseline assessment only and those who completed both a baseline and six-month assessment were not statistically significantly different. See Table 1.

Analysis.

Creating a matched control group. Youth who are referred to Wraparound may differ from youth referred for other types of mental health services in important ways, including acuity and complexity of mental health and other problems. For the purpose of designing a study that would

reduce selection bias and improve comparability between youth who were enrolled in Wraparound and youth who received other types of behavioral health services, we used propensity score matching to create a control group with similar demographic characteristics and baseline levels of impairment. Pair matching on a propensity score allowed us to estimate the average treatment effect for youth receiving Wraparound.²¹

Propensity to assignment to Wraparound was calculated via a logistic regression model, with age, race, sex, and the total number of actionable needs within each CANS domain at baseline (e.g., total number of “Child Risk Behaviors” scored either a 2 or 3 on the baseline assessment) used as covariates. We then matched using a nearest neighbor method, without replacement, setting caliper widths to 0.20 standard deviation of the logit of the propensity score²². Monte Carlo simulations have found that caliper methods without replacement have high performance relative to other matching methods, resulting in strong balance of baseline covariates and accurate estimates²³. Similar methods have found that caliper widths of 0.20 result in a reduction in bias and confidence intervals with approximately correct coverage rates compared to calipers of other widths²⁴.

To evaluate the success of the matching process, we followed the procedure described in Austin 2009²⁵, foregoing significance testing, and instead relying on standardized differences and visual inspections of variable balance. Matching was performed using MatchIt²⁶, and covariate balance was assessed in part using cobalt²⁷ in R version 3.5.2²⁸.

Figure 1 shows additional details about participant flow. Note that although the majority (89%) of eligible comparison youth were not included in analysis, in propensity-score matching methods the comparison group is meant only to approximate the counterfactual for exposed participants and not to estimate outcomes on the unexposed in the population. In our case, the

propensity-matched youth are not intended to be representative of all youth who were not enrolled in Wraparound.

Estimating the effect of Wraparound. We calculated the effect of Wraparound in three ways. First, we calculated the change in number of actionable needs (i.e., change from baseline to six-months in the number of items that were scored as a 2 or 3 across all items on the CANS), as well as within each of the seven CANS Domains. Items on the CANS are largely independent of one another, meaning that youth can improve on some items and get worse on others. This was true in our data. Using all baseline and six-month CANS available for this paper (n=6,679 CANS assessments), the correlation between pairs of CANS items ranged from -0.139 to 0.785. The 25th, 50th and 75th percentile correlations of item-pairs were 0.036, 0.084, and 0.154 respectively. Thus, most item-pairs are only weakly correlated, highlighting the independence of the CANS items in our study population. Therefore, we also compared youth receiving Wraparound and PSM-youth who did not receive Wraparound with respect to the number of actionable needs “resolved” (i.e., number of items actionable at baseline, that were no longer actionable at follow-up), and number of actionable needs “discovered” or “developed” (i.e., items not actionable at baseline, but actionable at follow-up), both across the CANS as a whole and within each of the seven Domains.

We used two statistical approaches to address study questions. Group differences in the average change in actionable needs from baseline to six-months were estimated using general linear regressions. Group differences in the number of actionable needs resolved and discovered were estimated using negative binomial regressions. Negative binomial regression can be used in the place of Poisson regression when data are over-dispersed²⁹. The PSM process creates balance

with respect to potential confounders, so no additional covariates were included in our models to compare Wraparound to the comparison group. Data were analyzed using R version 3.5.2²⁸. A two-sided p-value of .05 was set as the level of statistical significance.

RESULTS

Matching Diagnoses. The matching algorithm successfully matched 122 of the 129 Wraparound-enrolled youth to youth who received other services. This resulted in a final sample size of 122 Wraparound participants and 122 comparison participants. See Figure 1.

Balance of matching covariates was assessed primarily by calculating standard differences. Table 2 and Figure 2 describe covariate balance results. Twelve of the 13 covariates in the matched sample had a standard difference of 0.10 or less, a commonly-used threshold for assessing covariate balance²⁵. One covariate, the number of actionable needs at baseline in the “Behavioral and Emotional Needs” Domain, had a standardized difference just above the threshold, 0.103. Balance was also assessed by examining the similarity of boxplots for continuous variables between the exposed and unexposed groups. Overall, the variables were well-balanced between the groups.

Impact of Wraparound.

Tables 3-5 summarize the relative effect of Wraparound on the youth who received it, compared to the propensity-matched comparison group. Wraparound youth experienced a greater net change in actionable needs after six months (Mean Difference: -0.36, 95%CI: -2.03, +1.32), though confidence intervals were wide such that this difference was non-significant. Similarly, Wraparound youth experienced both a greater reduction in actionable needs (IRR: 1.32, 95%CI:

0.98, 1.79) and had new needs discovered more often than non-Wraparound youth (IRR: 1.23, 95%CI: 0.92, 1.65), though neither of these outcomes differed significantly between the two groups. All domain-level differences were also non-significant.

DISCUSSION

This paper used change on the CANS assessment to estimate the impact of Wraparound on youth behavioral, emotional, and functioning outcomes, relative to a propensity score matched comparison group of youth who received other kinds of services from the same provider. After six months, Wraparound-enrolled youth had more needs resolved and more new needs developed or discovered compared to the comparison group.

Baseline CANS scores for youth included in this study demonstrate the complexity of behavioral health needs that Wraparound is designed to target. On average, these youth began the Wraparound process with 20.5 actionable needs items on the CANS, indicating that over 20 conceptually independent areas were appropriate for intervention. For context, 20.5 needs represent 26% of the total items included in this analysis. As a comparison, in the same dataset, the average number of needs that youth had when beginning short-term residential stays at the same agency was 15.2, and the average number of needs for youth beginning Multisystemic Therapy was 18.4. A recent study by a Wraparound research team at the University of Washington found that, across 11 agencies' CANS data, the median number of actionable needs out of a common pool of 32 items across all agencies was 7 (22%)⁶, suggesting that the youth in this study are similar with regards to complexity of need to youth in Wraparound at other providers.

On average, Wraparound-enrolled youth had an actionable need in every domain, with the exception of the Cultural Factors domain. The caregivers of these youth also had needs; the average number of actionable needs within the Caregiver Resources and Needs domain was 2.52. We have no direct information about whether Wraparound youth were involved in multiple interventions and systems, but the acuity of need makes it likely that many were.

The work in the current study builds on the dissertation by Karpman, which also compared change in CANS scores for Wraparound youth and a PSM-matched comparison group. In that study, the addition of Wraparound was associated with a slower rate of change. In the previous study, functioning was measured using two net-measures of domain-level change, change in total number of actionable needs and change in the mean total score. Both of these measures obscure improvements made in cases where youth also worsen on other, independent items. For example, if a youth's baseline scores for the anger control and anxiety items are 3 and 0, respectively, and her six-month scores are 1 and 2 for the same items, the scoring indicates that she has resolved her need for clinical intervention on her anger control, but developed a need for intervention with her anxiety. Neither of the Karpman measures of change capture this type of change. In the prior study, the change in actionable needs would be 0 (1 actionable need at baseline and 1 actionable need at six months), and the change in mean score would also be 0 (3.00 at baseline and 3.00 at six months).

In addition to measuring the change in the number of actionable needs for the CANS as a whole and within each domain, the current study also measures the number of actionable needs that are resolved and the number that are newly developed after six months. The difference in net change in actionable needs between Wraparound-enrolled youth and the comparison group was modest (-0.36), but larger between-group differences emerged when comparing the number of

needs resolved (Incidence Rate Ratio: 1.32, 95% CI: 0.98, 1.79) and discovered (IRR: 1.23, 95% CI: 0.92, 1.65). Wraparound youth both resolved more needs and had a greater number of new needs at six months, though confidence intervals were wide, and the estimated differences were non-significant. These results highlight the need for future researchers who use the CANS as an outcome tool to think carefully about how change is measured and to consider reporting on improvement separately from worsening.

Although not statistically significant, in our study Wraparound youth improved on an estimated 32% more needs than the control group. In absolute terms, that difference amounted to a little more than one (1.13) additional need resolved after six months. The difference corresponds to a Cohen's *d*, a commonly used standardized measure of effect size, of 0.23. This is a "small" effect by the conventions initially set by Cohen.³⁰ However, the CANS is designed so that any change of one point is, by definition, clinically relevant. Clinical relevance is the foundation of the assessment's focus on "actionability." The resolution of an additional actionable need, which may result from a change of between 1 and 3 points (movement from a 2 or 3 to a 1 or 0), is therefore interpreted as a meaningful difference in the clinical course for a youth. The domain in which the estimated difference was largest was the Child Strengths domain (IRR: 1.43, 95% CI 0.99 – 2.33). Example items in this domain include child assets in the areas of Family Support, Resilience, and Community Life. Although items in this domain are written to describe a prosocial characteristic of the youth rather than a need, the items are scored similarly to the other items such that lower scores are preferable. Specifically, a score of 0 or 1 indicates the presence of a strength, while a 2 or 3 indicates the lack of a strength. Therefore, the equivalent change to "resolving" a need on the other domains (moving from a 2 or 3 to a 1 or 0) is "developing" a strength in the Child Strengths domain. Wraparound is an explicitly "strengths

based” process; a focus on strengths appears in the National Wraparound Initiative’s “10 Principles of Wraparound”³¹ and in the National Wraparound Implementation Center’s four Key Elements of Wraparound. Differences in the number of resolved needs were also relatively larger within the Behavioral and Emotional Needs (IRR: 1.36, 95%CI: 0.91 – 2.05), Child Risk Behaviors (1.31, 95%CI: 0.80 – 2.23), and Life Functioning (IRR: 1.31, 95% CI: 0.88 – 1.97) domains. These three domains cover the items most directly reflective of youth-level functioning. There was no estimated difference for the Cultural Factors domain (IRR: 1.00), and the Wraparound youth were estimated to resolve fewer needs in the Traumatic Adverse Childhood Experiences (TACE) domain.

Wraparound youth also had more new needs at six months than did the comparison group. While this may indicate a harmful effect of Wraparound on youth, we do not think that is the best explanation for this result. The appearance of new needs after six months may not always indicate that a child has developed a new problem. In some cases, a new need may appear due to the later disclosure or discovery of a need that was present but unacknowledged at the time of the baseline assessment. The domains in which new needs most commonly appeared were TACE and Caregiver Resources and Needs. The items in these domains cover sensitive topics, and it is reasonable to think that relevant information might not be readily disclosed at the time of intake but could emerge later as a Wraparound team built trust. The TACE domain includes several highly sensitive items (e.g., sexual abuse, physical abuse, parental criminal behavior, family violence) that a family may not be comfortable sharing at the time of treatment initiation. In the case of the Caregiver domain, caregivers seeking help for their children may be hesitant to disclose their own needs, seeing them as embarrassing or irrelevant. Relative to the providers of conventional behavioral health services, Wraparound care coordinators spend an

exceptional amount of time not only with enrolled youth, but also with the family and community of that youth. This greater level of integration into the family may make caregivers more willing to share sensitive information over time, or may allow Wraparound providers to discover needs of which the family was unaware.

LIMITATIONS

This study has several limitations. First, the study is observational, not experimental. Propensity score matching assures balance on measured and specified covariates, but it cannot assure that groups are equivalent with respect to other characteristics. While the matching in this study was successful, and the included covariates covered a breadth of demographic, emotional, and functional characteristics that might reasonably be associated with both Wraparound enrollment and clinical improvement, there are some baseline characteristics that were unmeasured and so not included. Relevant unmeasured variables included the type of insurance, the specific mental health diagnoses for each youth, and the youth's involvement with child serving systems such as the justice system and child welfare system. Of these, the last is likely the most important. Given the target population for the Wraparound process, it is possible that Wraparound youth were involved in a larger number of interventions and public systems than the matched youth, despite their comparable emotional and functional needs. The added complexity of multi-system involvement could affect the rate or magnitude of change in outcomes of Wraparound-enrolled youth.

Second, the CANS is a complicated measure to subject to quantitative analyses. There are many ways to score and interpret results. While this paper looks at average change in three ways (net change in actionable needs, the resolution of actionable needs, and the development of new

needs), there are other ways to measure change with the tool: the change in average score, the number of items to improve or worsen by a single point, calculating a reliable change index, and likely others. Based on the results of this paper, methods to measure change that best emphasize distinctions between the items are likely to show greater differences between groups. The resolution of an actionable need after six months of intervention is a high bar for change. Traditional psychometric assessments, which focus on individual symptoms, may better highlight the differential impact of Wraparound.

Third, we do not have information about the fidelity of Wraparound delivered. Wraparound is a complex process, requiring not only a skilled Wraparound care coordinator, but also a child-serving- and service-system amenable to care coordination. Previous research has suggested a relationship between low fidelity and poorer outcomes³². Although staff at the agency included in this paper all undergo a four-day Wraparound training and regular Wraparound-focused supervision, we do not have direct measures the fidelity or quality of practice for the treatment providers who worked with the study youth and families.

Finally, our outcomes are limited to change experienced by youth after six months. The timeframe for Wraparound is variable, but Wraparound is often a long process. Youth and families may be enrolled in the Wraparound process from anywhere from six months to over one year, though some families will stop services earlier for a variety of reasons. Thus, some Wraparound-enrolled youth did not complete six months of the process and so could not be included in this analysis, and many of the Wraparound youth who were included would likely stay enrolled in Wraparound for many more months. Although there were no major differences in the baseline characteristics of Wraparound youth who did and did not complete a six-month assessment (Table 1), the change experienced by these youth may have been different. That is,

youth for whom the process was not working may have disengaged too early to be included in this paper. Given the complexity of their needs, it is unlikely that many youth stopped Wraparound before six months after having most of their needs met. We do not know the relative effectiveness of Wraparound after the six-month period. Future research should focus on the longitudinal impact of Wraparound, both during and after Wraparound enrollment.

Conclusion

The current study contributes to the growing effectiveness research base on the behavioral, emotional, and functional impacts of Wraparound in several ways. First, it uses novel methodology that can be replicated in future studies. In the absence of randomized experiments, propensity-scores provide a powerful and intuitive method to reduce the bias inherent in observational studies. Future research using PSM can use results of this study to estimate power. Second, it demonstrates ways in which service systems that use the CANS can maximize the utility of the CANS, by considering the number and pattern of items that were resolved versus emerged from baseline to follow-up, rather than looking at total scores and net change.

Finally, although the estimated effects of Wraparound in this study only approached significance, between-group differences showed a moderate effect size and were clinically meaningful as per conventions for interpreting results from the CANS. This, combined with the observed pattern of types of needs that were resolved versus emerged over 6 months, provides modest, but meaningful, evidence in favor of Wraparound's effectiveness as implemented by this provider. Given that there are many hundreds of Wraparound provider organizations nationwide, the study demonstrates the potential these Wraparound providers have for contributing to research base through consistent outcomes measurement. By using standardized measures such

as CANS consistently and over longer periods – including at the time of treatment discharge – in combination with methods such as PSM, consortia of provider organizations could pool their data and/or results, establish benchmarks for functional improvement, and greatly aid the cause of understanding the potential for positive impact of Wraparound and other community-based service models.

FIGURES AND TABLES

Fig 1. Participant flow.

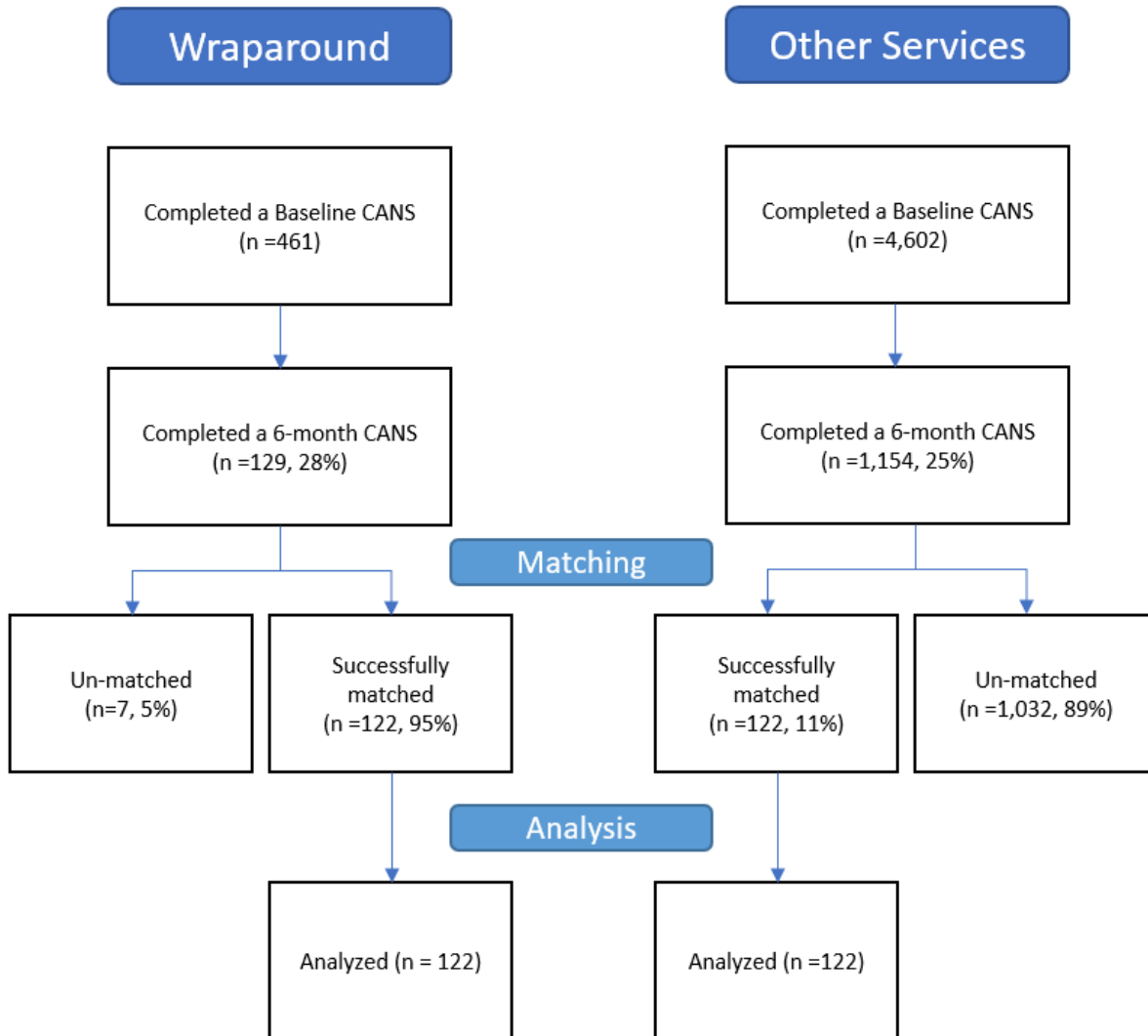


Figure 2. Propensity Score Matching Covariate Balance

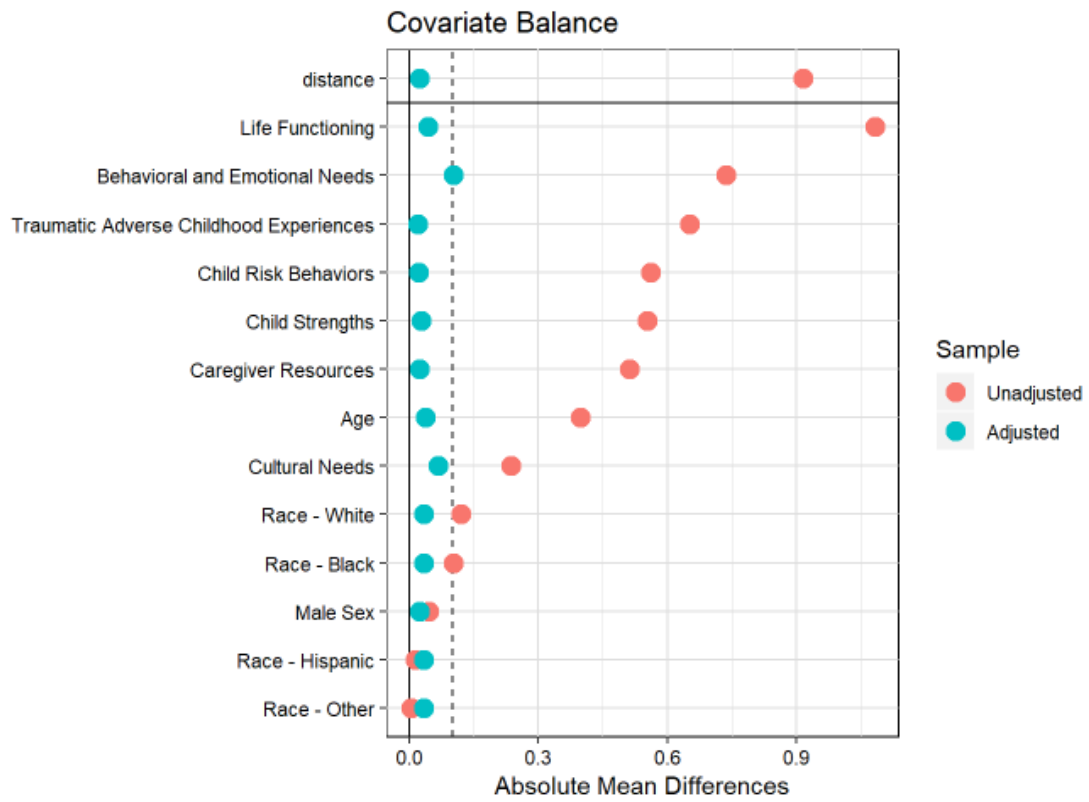


Table 1. Average baseline characteristics of Wraparound-enrolled youth who completed a baseline CANS assessment and youth who completed both a baseline and six-month assessment.

Covariate	Baseline and 6m	Baseline Only	<i>P</i>*
Behavioral and Emotional Needs	3.25	2.95	0.06
Caregiver Resources and Needs	2.61	2.36	0.31
Life Functioning	4.53	4.07	0.04
Child Risk Behaviors	1.19	1.48	0.02
Cultural	0.39	0.29	0.21
Traumatic Adverse Childhood Experiences	2.60	2.35	0.23
Child Strengths	6.48	6.21	0.39
Male Sex	0.56	0.51	0.35
Age	13.73	13.83	0.75
Race			
White	0.28	0.27	0.88
Black	0.22	0.22	0.97
Hispanic	0.38	0.38	0.93
Other	0.12	0.13	0.91

* two-tailed t-test

Table 2: Balance of matching covariates

Covariate	Wraparound Mean	Non- Wraparound Mean	Cohen's d
Behavioral and Emotional Needs	3.20	3.05	0.103
Caregiver Resources and Needs	2.52	2.46	0.022
Life Functioning	4.34	4.25	0.046
Child Risk Behaviors	1.15	1.18	0.021
Cultural Factors	0.37	0.33	0.073
Traumatic Adverse Childhood Experiences	2.58	2.54	0.019
Child Strengths	6.34	6.26	0.027
Male Sex	0.55	0.53	0.049
Age	13.64	13.53	0.037
Race			
White	0.27	0.3	0.072
Black	0.23	0.2	0.079
Hispanic	0.38	0.35	0.067
Other	0.12	0.15	0.094

Table 3. Net change of actionable needs after six-months for Wraparound-enrolled youth and PSM-matched youth¹¹.

	Wraparound Mean	Non-Wraparound Mean	Mean Difference	95% CI	p*
Total CANS	-1.46	-1.11	-0.36	-2.03, 1.32	0.68
CANS Domains					
Behavioral and Emotional Needs	-0.59	-0.36	-0.24	-0.62, 0.14	0.23
Caregiver Resources and Needs	0.49	0.39	0.10	-0.46, 0.66	0.73
Child Strengths	-0.85	-0.45	-0.41	-1.01, 0.20	0.19
Life Domain	-0.50	-0.41	-0.10	-0.57, 0.38	0.69
Child Risks and Behaviors	-0.18	-0.18	0.00	-0.21, 0.21	0.99
Traumatic Adverse Childhood Experiences	0.16	-0.07	0.23	-0.01, 0.47	0.06
Cultural Factors	0.02	-0.04	0.06	-0.02, 0.14	0.16

* General linear regression

¹ Note: Negative change scores indicate fewer actionable needs at six months, compared to baseline.

Table 4. Number of actionable needs resolved after six-months for Wraparound-enrolled youth and PSM-matched youth.

	Wraparound Mean	Non-Wraparound Mean	IRR	95% CI	P*
Total CANS	4.63	3.50	1.32	0.98, 1.79	0.07
CANS Domains					
Behavioral and Emotional Needs	0.95	0.70	1.36	0.91, 2.05	0.14
Caregiver Resources and Needs	0.57	0.50	1.15	0.60, 2.19	0.67
Child Strengths	1.43	0.94	1.52	0.99, 2.33	0.06
Life Domain	1.09	0.83	1.31	0.88, 1.97	0.19
Child Risks and Behaviors	0.39	0.29	1.33	0.80, 2.23	0.27
Traumatic Adverse Childhood Experiences	0.16	0.21	0.77	0.37, 1.58	0.48
Cultural Factors	0.03	0.03	1.00	0.24, 4.23	0.99

* Negative binomial regression

Table 5. Number of new actionable needs discovered or developed after six-months for Wraparound-enrolled youth and PSM-matched youth.

	Wraparou nd Mean	Non- Wraparou nd Mean	IRR	95% CI	P*
Total CANS	3.28	2.67	1.23	0.92, 1.65	0.16
CANS Domains					
Behavioral and Emotional Needs	0.33	0.36	0.91	0.57, 1.45	0.69
Caregiver Resources and Needs	1.11	0.94	1.17	0.80, 1.73	0.42
Child Strengths	0.20	0.12	1.11	0.67, 1.85	0.69
Life Domain	0.59	0.45	1.31	0.86, 2.00	0.21
Child Risks and Behaviors	0.34	0.20	1.67	0.78, 3.66	0.19
Traumatic Adverse Childhood Experiences	0.66	0.59	1.75	0.89, 3.48	0.11
Cultural Factors	0.07	0.01	8.00	1.37, 152.25	0.06

* Negative binomial regression

References

1. O'Connell ME, Boat TF, Warner KE, National Research Council (U.S.). Board on Children and Families Y, Institute of Medicine (U.S.). Committee on Prevention of Mental Disorders and Substance Abuse Among Children and Young Adults: Research Advances and Promising Interventions. Institute of Medicine U.S. . Committee on Prevention of Mental Disorders and Y. *Preventing Mental, Emotional, and Behavioral Disorders among Young People : Progress and Possibilities*. (O'Connell ME, Boat TF, Warner KE, National Research Council (U.S.). Board on Children and Families Y, eds.). Washington, D.C.: Washington, D.C. : National Academies Press; 2009.
2. Law P, Act A. the United States of America in Congress assembled , ADMINISTRATION AND INSTITUTES Subtitle A-Administration. 1992.
3. Williams NJ, Scott L, Aarons GA. Prevalence of Serious Emotional Disturbance Among U.S. Children: A Meta-Analysis. *Psychiatr Serv*. 2017;69(1):32-40. doi:10.1176/appi.ps.201700145
4. Pires SA, Grimes KE, Allen KD, Gilmer T, Mahadevan RM. Faces of Medicaid: Examining Children's Behavioral Health Service Use. 2013;(December).
5. Bruns EJ, Walker JS, Zabel M, et al. Intervening in the lives of youth with complex behavioral health challenges and their families: the role of the wraparound process. *Am J Community Psychol*. 2010;46(3-4):314-331. doi:10.1007/s10464-010-9346-5
6. Schurer Coldiron J, Bruns EJ, Quick H. A Comprehensive Review of Wraparound Care Coordination Research, 1986–2014. *J Child Fam Stud*. 2017;26(5):1245-1265. doi:10.1007/s10826-016-0639-7
7. Clark H, Lee B, Prange M, McDonald B. Children lost within the foster care system: Can wraparound service strategies improve placement outcomes? *J Child Fam Stud*. 1996;5(1):39-54.
8. Deaner A. A comparative evaluation of wrap-around services and partial hospitalization for preschool children exhibiting severe behavior problems. Rosner S, ed. 1998.
9. Ferguson CM, Michael C. Copyright 2004 by. 1997.
10. Carney MM, Buttell F. Reducing juvenile recidivism: Evaluating the wraparound services model. *Res Soc Work Pract*. 2003;13(5):551-568. doi:10.1177/1049731503253364
11. Bruns EJ, Pullmann MD, Sather A, Ramey M, Brinson RD. Effectiveness of Wraparound Versus Case Management for Children and Adolescents: Results of a Randomized Study. *Adm Policy Ment Heal Ment Heal Serv Res*. 2014;42(3):309-322. doi:10.1007/s10488-014-0571-3
12. Aboutanos MB, Jordan A, Cohen R, et al. Brief violence interventions with community case management services are effective for high-risk trauma patients. *J Trauma - Inj Infect Crit Care*. 2011;71(1):228-237. doi:10.1097/TA.0b013e31821e0c86
13. Csokasy J. An effectiveness study of wraparound care. Douglass B, ed. 1997.
14. Mears SL, Yaffe J, Harris NJ. Evaluation of Wraparound Services for Severely Emotionally Disturbed Youths. *Res Soc Work Pract*. 2009;19(6):678-685. doi:10.1177/1049731508329385
15. Karpman H. Is Everything Coming up Rosie? What is the Impact of a Redesigned Medicaid Funded Mental Health System on Child Mental Health, Caregiver Needs and Strengths, and Youth Risk Factors for Substance Abuse? Hodgkin D, Perloff J, Reif S, Simons J, eds. 2013.

16. Grimes KE, Schulz MF, Cohen SA, Mullin BO, Lehar SE, Tien S. Pursuing cost-effectiveness in mental health service delivery for youth with complex needs. *J Ment Health Policy Econ*. 2011;14(2):73.
17. Lyons JS. *Communitometrics: A Communication Theory of Measurement in Human Service Settings*. New York, NY: New York, NY: Springer US; 2009. doi:10.1007/978-0-387-92822-7
18. Anderson RL, Lyons JS, Giles DM, Price JA, Estle G. Reliability of the Child and Adolescent Needs and Strengths-Mental Health (CANS-MH) scale. *J Child Fam Stud*. 2003;12(3):279-289. doi:10.1023/A:1023935726541
19. Coldiron JS, Hensley S, Parigoris R, Becker JR, Parker EM, Bruns EJ. Creating Cross-site CANS Benchmarks: Profiles of Baseline Needs and 6-month Change in a Large Multi-site Sample of Wraparound-Enrolled Youth. In: *30th Annual Research & Policy Conference on Child, Adolescent, and Young Adult Behavioral Health*. Tampa, FL; 2017.
20. Bruns EJ, Schurer Coldiron J, Hensley S. Using administrative CANS data for benchmarking and outcomes monitoring in state-wide Wraparound initiatives. In: *29th Annual Research & Policy Conference on Child, Adolescent, and Young Adult Behavioral Health*. Tampa, FL; 2016.
21. Imbens GW. Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review. *Rev Econ Stat*. 2004;86(1):4-29. doi:10.1162/003465304323023651
22. Rosenbaum PR, Rubin DB. Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score Published by : Taylor & Francis , Ltd . on behalf of the American Statistical Association Stable URL : <http://www.jstor.org/stable/2683903> Your. *Am Stat*. 1985;39(1):33-38.
23. Austin PC. A comparison of 12 algorithms for matching on the propensity score. *Stat Med*. 2014;33(6):1057-1069. doi:10.1002/sim.6004
24. Austin PC. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharm Stat*. 2011;10(2):150-161. doi:10.1002/pst.433
25. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med*. 2009;28(25):3083-3107. doi:10.1002/sim.3697
26. Ho DE, King G, Stuart EA, Imai K. MatchIt : Nonparametric Preprocessing for Parametric Causal Inference. *J Stat Softw*. 2011;42(8):1-28. doi:10.18637/jss.v042.i08
27. Greifer N. cobalt: Covariate Balance Tables and Plots. R package version 3.6.1. 2019. <https://cran.r-project.org/package=cobalt>.
28. Team RC. R: A language and environment for statistical computing. 2018. <https://www.r-project.org/>.
29. Stamey CJD, Beavers D, Kattan MW. Encyclopedia of Medical Decision Making Poisson and Negative Binomial Regression. 2019:886-889.
30. Cohen J. Statistical power analysis for the behavioral sciences. 1977.
31. Bruns EJ, Walker JS. Ten Principles of the Wraparound process. *Resour Guid to Wraparound*. 2010;(April 2010):Chapter 2.1.
32. Effland VS, Walton BA, McIntyre JS. Connecting the Dots: Stages of Implementation, Wraparound Fidelity and Youth Outcomes. *J Child Fam Stud*. 2011;20(6):736-746. doi:10.1007/s10826-011-9541-5

**Appendix A: Items from the Child and Adolescent Needs and Strengths (CANS)
Assessment, organized by domain**

Behavioral / Emotional Needs	Life Functioning
Anger Control	Decision Making
Anxiety	Developmental Intellectual
Conduct	Family Functioning
Depression	Independent Living
Eating Disturbance	Legal
Impulsivity Hyperactivity	Living Situation
Oppositional	Medical
Psychosis	Physical
Somatization	Recreational
Substance Use	School
Caregiver Needs	Sexual Development
Developmental	Sleep
Family Stress	Social Functioning
Involvement with Care	Youth Risk Behaviors
Knowledge	Criminal Behavior
Legal	Danger to Others
Mental Health	Fire Setting
Organization	High Risk Sexual Activity
Physical	Intentional Misbehavior
Residential Stability	Other Self Harm
Safety	Runaway
Social Resources	Self Injurious Behavior
Substance Use	Sexual Aggression
Supervision	Sexual Exploitation
Cultural Factors	Suicide Risk
Cultural Differences within the Family	Traumatic Adverse Childhood Experiences
Cultural Identity	Disruption in Caregiving
Discrimination Bias	Emotional Abuse
Language	Medical Trauma
Traditions and Rituals	Natural Disaster
Individual Strengths	Neglect
Community Life	Parental Criminal Behavior
Coping and Savoring Skills	Physical Abuse
Educational Setting	School Violence
Family Strengths	Sexual Abuse
Interpersonal	Terrorism Affected
Natural Supports	War Affected
Optimism	Witness to Community Violence
Relationship Permanence	Witness to Family Violence
Resiliency	Witness to Criminal Acts
Resourcefulness	
Spiritual Religious	
Talents Interests	
Vocational	