



Published in final edited form as:

Child Youth Serv Rev. 2008 ; 30(7): 746–757.

Comparing Outcomes for Youth in Treatment Foster Care and Family-style Group Care

Bethany R. Lee, Ph.D. and

Assistant Professor, School of Social Work, University of Maryland Baltimore, 525 W. Redwood St., Baltimore, MD 21201, blee@ssw.umaryland.edu, 410/706-3809

Ron Thompson, Ph.D.

*Director, National Research Institute, Girls and Boys Town Boys Town, NE 68010
thompsonr@girlsandboystown.org, 402/498-3102*

Abstract

Group care programs are often criticized for producing poor outcomes, especially in light of community-based alternatives like treatment foster care that have a stronger evidence base. In this study, data from Girls and Boys Town were used to compare outcomes of youth in treatment foster care (n=112) and group care (n=716) using propensity score matching, a method that can minimize selection bias in nonrandomized designs. Eighteen background covariates were used to develop propensity scores for the likelihood of receiving treatment foster care rather than group care. Several matching methods generated balanced samples on which the outcomes were compared. Results found that group care youth were more likely to be favorably discharged, more likely to return home, and less likely to experience subsequent placement in the first six months after discharge. Legal involvement and residing in a home-like environment at follow-up did not differ. Positive outcomes for group care youth suggest that family-style group care programs may promote effectiveness.

Keywords

treatment foster care; family-style group care; propensity score matching

Prevailing wisdom in systems of care for youth promotes a philosophy of least restrictive and community-based care (Stroul & Friedman, 1986). A growing body of evidence suggests that specialized foster care settings like treatment foster care (Chamberlain & Reid, 1991; Fisher & Chamberlain, 2000) are more effective and cost-efficient than group care (Chamberlain & Reid, 1998; Eddy & Chamberlain, 2000; U.S. Department of Health and Human Services [USDHHS], 1999). In a continuum of care, group care often serves youth whose needs surpass the capacity of traditional family foster care, but are not acute enough to warrant inpatient hospitalization (CWLA, 2004). With the emergence of treatment foster care as a possible alternative to group placements, questions about the continued broad use of group care are increasing (Barth, 2005).

Despite widespread questions about youth safety during care (Barth, 2002) and equivocal outcomes following care (Hair, 2005; Wells, 1991), group care programs are prevalent. In this article, the term group care will be used interchangeably with residential care to label programs

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

that provide services 24-hours a day to groups of dependent and/or emotionally disturbed youth (Pecora, Whittaker, Maluccio, Barth & Plotnick, 1992). Of the over half a million youth in out-of-home placements, almost 1 in 5 live in group care settings (USDHHS, 2003). Among older youth, the proportion is much greater (McMillen et al., 2005). These youth represent some of the most troubled and troubling clients of public child-serving systems, including child welfare, juvenile justice and mental health (Coen, Libby, Price & Silverman, 2003).

Youth with externalizing disorders are often placed in specialized settings like group homes, where they live with other youth who have similar problem behaviors. In a nationally representative sample of youth who had spent about a year in the child welfare system, 55% of youth in group care scored in the clinical range of the CBCL delinquency subscale, compared to 48% in non-kin foster care (USDHHS, 2001a). Rates of conduct disorder or oppositional defiant disorder diagnoses for youth entering group care have been reported to be as high as 75% (Handwerk, Field & Friman, 2001) and more than twice the rate for youth in foster family settings (McMillen et al., 2005).

Treatment foster care (TFC) has been developed as a less-restrictive community-based alternative to group care. Although TFC is often considered as a substitute for group care, the population of youths served in TFC and group care show some differences on average. Studies that have descriptively compared TFC and group care have found that TFC youths are less disturbed (Berrick, Courtney & Barth, 1993), less likely to be taking psychotropic medications (Breland-Noble et al., 2004), and more likely to receive community-based supplementary services (Breland-Noble, Farmer, Dubs, Potter & Burns, 2005) than group care youth.

However, it is not clear whether aggregate differences in youth populations served by TFC and group care are intentional or accidental. While policy and practice trends suggest that less troubled youth should be placed in less-restrictive settings, not enough evidence is available to anticipate which youth are best served in which environments. There are no guidelines or admissions criteria for determining who should receive group placement (Whittaker, 2000). Without specified protocols for placement decision-making, it is difficult to understand how a youth ends up in a treatment foster care rather than group care placement.

Previous studies comparing one model of treatment foster care (TFC) with group care found that TFC leads to better delinquency and adjustment outcomes for youth. However, several problems have been identified in these comparative studies, including poor response rate, small sample size, an emphasis on process rather than outcome measures, and a lack of description regarding differences in treatment philosophy (Bates, English & Kouidou-Giles, 1997; Curtis, Alexander & Lunghofer, 2001; Drais-Parillo, 2004). This study minimized the limitations present in earlier efforts by utilizing a larger sample size, assessing outcomes at discharge and follow-up (with over 75% response rate), and comparing programs with the same treatment orientation.

The purpose of this paper is to compare delinquency and adjustment outcomes of youth who received treatment foster care with youth who received a family-style group care program. While this study was observational rather than experimental, propensity score matching was used to balance the selection bias present in non-randomized designs. Based on the promising results of TFC in earlier studies, it was hypothesized that youth in treatment foster care would be more likely to experience positive outcomes in delinquency and adjustment domains compared to their matched peers who received group care.

Evidence for Treatment Foster Care's Effectiveness

Currently, the only established evidence-based intervention for out-of-home placement is a structured treatment foster care model. Multidimensional Treatment Foster Care (MTFC)

emphasizes close adult supervision, fair and consistent limits, predictable consequences, supportive adult relationships and limited exposure to deviant peers by placing youth individually in a specially trained foster family that received additional supports, including daily phone calls and 24-hour response availability (Fisher & Chamberlain, 2000). MTFC has been recognized as a model program in the 1999 Surgeon General Report on Mental Health (USDHHS, 1999), the 2001 Surgeon General Report on Youth Violence (USDHHS, 2001b), and labeled a “Blueprint” program by the Center for the Study and Prevention of Violence (Chamberlain & Mihalic, 1998).

In studies comparing youth randomized into either MTFC or diverse group care programs, Chamberlain and colleagues found that MTFC youth spent less time in locked settings, had fewer criminal referrals (Chamberlain, & Reid, 1998; Leve, Chamberlain & Reid, 2005) and had fewer delinquent peers at follow-up (Leve & Chamberlain, 2005). Males in MTFC were more likely to complete treatment and return home to family compared to males in group care (Chamberlain & Reid, 1998).

The randomized experiment comparing youth in MTFC and group care conducted by Chamberlain and colleagues made a strong argument in favor of alternatives to group care. However, the within-group differences among the group care programs may have impacted results. In the Oregon study, youth randomized into group care received diverse treatment with the population size of the group care settings ranging from 2 to 52 youths (Leve & Chamberlain, 2005). These 19 different group care programs applied varied theoretical approaches (Chamberlain & Reid, 1998). Most of the programs were described as having a Positive Peer Culture orientation, which emphasize peers as important agents of change (Vorrath & Brendtro, 1985) and has little evidence for its effectiveness. In a study comparing staff assumptions and program practices, many differences were found between these programs and the MTFC model (Chamberlain, Ray & Moore, 1996). Because of these confounding factors, other researchers have identified the need for a comparison of group care and treatment foster care programs that differ in fewer respects (Osgood & Bridell, 2006; Dodge & Sherrill, 2006). This study achieves that goal by comparing group care provided under one longstanding and promising approach—the Teaching Family Model—and by using a rigorous statistical technique to estimate what would have happened if children had been randomly assigned to group care and treatment foster care.

Evidence for Family-style Group Care’s Effectiveness

Although much of the research on group care programs provides weak evidence for its effectiveness (cf. Frensch & Cameron, 2002; Hair, 2005), one model of group care, the Teaching-Family model, is a notable exception. Originally developed at Achievement Place in the 1970s (Wolf et al., 1976), the Teaching-Family model is a behaviorally-oriented approach that integrates family-style living, youth self-government and a token economy for teaching social skills (Friman, 2000; Larzelere, Daly, Davis, Chmelka & Handwerk, 2004). In Teaching-Family group homes, 6–8 youth live in a large home with live-in houseparents who provide consistent supervision and maintain a family-like environment through relationship-building. The low youth-to-staff ratios and high rates of positive reinforcement for socially desirable behaviors are components of the model that have been found to be effective (Friman, Jones, Smith, Daly, & Larzelere, 1997; Friman, Toner, Soper, Sinclair & Shanahan, 1996).

While no residential care model has met criteria as an evidence-based practice, over 100 studies have examined the Teaching-Family model, with most identifying promising results (Fixsen & Blase, 2002). In a meta-analysis of interventions for delinquency, the Teaching-Family model was found to be one of the most effective programs for decreasing re-arrest (Lipsey & Wilson, 1998). Compared to other group home models, youth from Teaching-Family programs

were three times less likely to experience subsequent delinquency charges in the first few months following placement (Youngbauer, 1997). Additional outcome domains where gains have been identified include educational achievement and positive relationships with adults and peers (Thompson et al., 1996).

As this brief overview suggests, Teaching-Family group care is quite different from the eclectic group care programs previously used in comparison studies with treatment foster care. Although the work of Chamberlain and colleagues has shown that a TFC model can outperform typical group care programs, this is the first study to compare a well-established and promising group care intervention with a treatment foster care model. In this study, both the group care and TFC programs have adopted the Teaching-Family approach. By holding the treatment model constant, differences in performance across the interventions can be more cleanly attributed to differences in the modality of the intervention (i.e. group or family care)..

Propensity Score Matching

Although random assignment to treatment group is the best approach for comparing two interventions like TFC and group care, randomization is not always possible, practical or ethical in real world practice. Guo and colleagues (2006) asserted that the costs of true experiments often impede the sample size, threatening statistical conclusion validity. Because of the challenges inherent in conducting randomized experiments in many social science research settings, methods like propensity score matching can be used to strengthen quasi-experimental designs. Quasi-experiments generally incur fewer costs and with propensity score matching, selection bias introduced due to lack of randomization, may be adequately diminished. In their study of a school-based intervention, results comparing propensity score methods and randomization showed differences only in tenths or hundredths of units (Guo, Rose & Bowen, 2006).

In propensity score matching, a propensity score is computed for each individual in the sample and reflects his/her probability to be assigned to a treatment group given observed background covariates (Rosenbaum & Rubin, 1983). While the propensity score can be used to balance dozens or even hundreds of observed covariates, it is itself a single variable that can then be easily used for matching participants in each treatment group (Guo, Barth & Gibbons, 2005). By comparing individuals from different treatment groups with the same or similar propensity scores, causal effects of the treatments can be estimated (D'Agostino, 1998; Winship & Morgan, 1999).

Of course, propensity scores can only match (i.e. balance) covariates that are observed. Unmeasured differences between individuals can not be directly addressed in non-randomized designs. However, post-hoc sensitivity analyses can be conducted to assess the magnitude of hidden bias that would need to exist in order to substantively alter findings (Braitman & Rosenbaum, 2002; Leow, Marcus, Zanutto & Boruch, 2004).

Applications of propensity score matching technique can be found in a variety of disciplines, including medicine (Seeger, Williams & Walker, 2005; Mitra & Indurkha, 2005; Iezzoni, 2003), education (Leow, Marcus, Zanutto & Boruch, 2004), public policy (Dahl & Lorentzen, 2005) and social work (Fraser, 2004; Sosin, 2002; El-bassel, Gilbert, Wu, Go & Hill, 2005). A growing interest in the utilization of this method is evident in child welfare studies (Guo, Barth & Gibbons, 2005; DeSena et al., 2005). Given the preponderance of nonrandom assignment to service use and types as well as the increased emphasis on outcomes, data from youth in public sectors of care are often well-suited for this method of analysis.

Methods

Agency description

This longitudinal study analyzed administrative data of youth in Girls and Boys Town programs. Founded in 1917, Girls and Boys Town has a long history of providing services to youth in out-of-home care. In addition to the residential program on the “home” campus in Boys Town, Nebraska, Girls and Boys Town USA has satellite agencies around the country that offer a variety of youth services, including treatment foster care programs.

Both the residential care and treatment foster care programs follow the Teaching-Family model. All youth care providers (including treatment foster parents) complete training on model implementation. Supervision is provided to foster parents and direct care staff to promote model fidelity.

Girls and Boys Town (GBT) was an ideal agency for this study for several reasons. First, the population of youth in residential care at GBT is large; there are 75 group homes on their main campus, totaling a population of more than 450 youth. Large sample sizes better accommodate advanced quantitative analyses like propensity score matching. Second, unlike many residential or treatment foster care programs, GBT provides a standardized intervention, in which fidelity to the Teaching-Family model is closely monitored. Third, both the group care and foster care programs utilize the same model, a key feature not found in other comparative studies. Data collection and management at GBT is more sophisticated than most youth-serving organizations. At intake, an extensive social history for the youth and family are recorded based on youth and family self-report as well as archival records, when available. At placement termination, status of discharge (favorable or unfavorable) as well as discharge setting is collected. Following placement, six-months post-discharge telephone interviews are conducted to track youth outcomes related to living situation and community functioning.

A de-identified subset of the Girls and Boys Town administrative database was provided to the investigator for this study. Study methods were approved by the Institutional Review Board (IRB) at Washington University in St. Louis as well as the Girls and Boys Town IRB. All analyses were conducted using Stata v9.

Sample description

All youth who were in placement at least 30 days and exited from residential care at Girls and Boys Town’s Home campus or Girls and Boys Town USA treatment foster care sites between June 1, 2002 and December 31, 2005 were included in the study sample. A 30-day minimum stay was selected because only youth who remained in placement at least 30 days were followed post-placement. This inclusion criteria eliminated 30 group care youth (3%) and 17 TFC youth (7.6%) whose length of stay was less than 30 days. To increase comparability between the two samples, only youth who were at least 8 years old at intake were included (excluding an additional 80 TFC youth). Youth with incomplete intake data or unavailable outcome measures were further excluded in these analyses.

Youth in treatment foster care received services in six geographic locations throughout the country (North Florida, Louisiana, Washington, DC, Philadelphia, Western Iowa/Nebraska, and New England). These sites contracted primarily with local child welfare jurisdictions. Although site-specific differences were not the focus of this study, preliminary chi-square analyses were conducted to assess whether any of the individual sites had vastly differential performance on youth outcomes. No significant differences were found across the TFC sites despite their disparate geographic locations. Youth in group care all received services at the Girls and Boys Town home campus in Omaha, NE.

Measures

Eighteen variables collected at intake were used to develop propensity scores. The selection of these items was based on the limited empirical evidence that identifies factors that predict placement type (TFC vs. group care) or are associated with youth outcomes (Collins, Schwartz, & Epstein, 2001; Drais-Parillo, 2004; Gorske, Srebalus & Walls, 2003; Piotrowski & Baker, 2004; Ryan, 2006; Sunseri, 2001; Teare et al., 1999). In addition to basic demographics (age, gender, race), the population density of the county of origin was also considered. Due to concerns about geographic variation in the use of residential placement (McMillen et al., 2005), it was hypothesized that urbanicity may impact likelihood of group placement. To classify county populations, the 2003 Rural/Urban Continuum Codes used by the Department of Health and Human Services for HealthyPeople 2010 was adapted. Each youth's county of origin was categorized as urban (population over 1 million), small city (250,000-1 million), or nonmetropolitan (less than 250,000).

Youth problem behavior history was also incorporated in the development of propensity scores. A nine-item Behavior Problem Index was used as a summative scale that assessed youth involvement in the following behaviors: substance use, physically assaultive toward adults, physically assaultive toward peers, possession of a weapon, running away, school behavior problems, verbally aggressive toward others, vandalism, or other criminal activity. Further evidence of previous known delinquent behavior was measured with items that included previous placement in a detention center, legal status as delinquent and formal system referral for placement.

A youth's maltreatment history was assessed at intake through youth or caregiver report or formal record of reported child abuse or neglect. This included previous physical abuse, sexual abuse or neglect. Family problems were also considered in developing propensity scores for matching youth. Specifically, the presence of mental health need, corrections involvement, substance use, and domestic violence within the family of origin were assessed.

Five outcomes related to delinquency and adjustment were assessed at discharge or the 6-month follow-up telephone interview. At discharge, the status of each youth was rated by a clinical specialist involved in his or her care. This Masters-level professional had frequent contact with the youth, family, and substitute caregivers throughout placement. Youth were assessed as receiving a favorable or unfavorable discharge based on the reason for discharge and completion of treatment goals. The outcome "return home" was measured as youth who were discharged to parents, relatives, or adoptive parents.

Follow-up data was collected for 78% of the youth in residential care (n=558) and 75% of treatment foster care youth (n=86). To measure criminal involvement since placement, respondents were asked about any arrests, convictions, probation or jail sentences since leaving the program. An endorsement of any of these items was considered a negative outcome. Respondents were asked about whether they had experienced any subsequent formal placements since discharge. Finally, the level of institutionalization of the youth's current living situation was assessed. Youth who resided in a community-based family home setting (i.e. foster home, home of a parent, relative or friend) were considered to be living in a homelike setting.

Statistical methods

Propensity score matching was used to compare outcomes for treatment foster youth matched to the residential care sample. Initial differences in covariates across groups were assessed to establish a baseline of standardized bias. Propensity scores were generated using logistic regression, modeling the likelihood for receiving treatment foster care. Placement type was

predicted by the 18 background covariates described previously. These covariates were expected to differ between groups and be significantly associated with the outcomes. The logit of the predicted probability output by this model was the propensity score.

Using PSMATCH2 (Leuven & Sianesi, 2003), four matching strategies were employed: three variations of nearest neighbor matching (one-to-one without replacement, one-to-one with replacement with a caliper, and one-to-six with replacement with caliper), and Gaussian kernel matching. In nearest neighbor matching, cases in each placement group are randomly ordered and then matched with the case from the other treatment condition with the nearest propensity score. One-to-one matching resulted in matched pairs while one-to-six matching allowed up to six group care youth to be matched to each TFC youth. By allowing replacement, a group care youth can be matched to more than one TFC youth. A caliper limits the distance of how far a “nearest neighbor” could be. In this study, a caliper of one-quarter of the standard deviation of the propensity scores (Rosenbaum and Rubin, 1986) was used.

In kernel matching, weights account for differences in match quality. The counterfactual outcome is computed for each member of the treatment group (i.e., TFC) as a kernel-weighted average of outcomes from all controls (i.e. group care). Larger weights are awarded to cases with closely matched propensity scores than more distal cases. Gaussian kernel matching with a bandwidth of .06 was used. Gaussian kernels have been shown to be the best-performing and a .06 bandwidth is most common (Essama-Nssah, 2006).

Matching algorithms often involve trade-offs between the quality (i.e. exactness) of the matches or the quantity of matches (i.e. number of treated cases in sample who are matched). For example, in this study, the number of group care youth was about six times larger than the TFC sample size. Matching only one group care youth to each TFC youth did not capitalize on the large sample of group care youth and left many group care youth unmatched. In contrast, kernel matching uses all cases without censoring poor matches. By utilizing multiple methods and jointly considering their results, the robustness of findings can be better assessed.

Following matching, the rate of achievement of the outcome of interest was compared across groups. Boot-strapped standard errors and 95% confidence intervals were used to assess whether differences in outcome achievement were statistically significant. Although it is impossible to test whether relevant covariates were missing from the model, new developments in Stata can assess the robustness of the findings to unmeasured covariates. Sensitivity analyses using the Stata ado-file mhbounds was conducted. This process offered guidelines on how large the hidden bias due to missing variables would need to be in order to substantively change study results.

Results

Prior to interpreting outcome results, the matched samples produced by each of the matching methods were evaluated to detect remaining selection bias between groups. Table 1 displays basic descriptive data by treatment group and any significant between group differences before and after matching. Note that actual chi-square and t-values are not provided post-match as these numbers differed for each combination of matching method and outcome.

As evident in Table 1, the two groups differed significantly on 14 of 18 covariates prior to matching. After matching, 13 of the 18 covariates were balanced across all four matching strategies. Of the remaining five covariates, three covariates showed imbalance following matching for one of the four matching strategies and two other covariates showed imbalance with two matching methods.

Covariate imbalance was interpreted for each of the five outcomes and these differences are described below. For favorable discharge status and returning home, group care youth were more likely to have family problems that included substance abuse (83% vs. 67%) or domestic violence (53% vs. 40%). This imbalance would suggest that group care youth may be at a disadvantage in achieving desirable outcomes. In looking at subsequent formal placement and homelike setting at follow-up, two matching strategies found TFC youth having significantly fewer prior placements than the matched group care sample (3.2 vs. 5.2). Since multiple placement changes have been found to be associated with negative outcomes, this imbalance suggests that the group care youths are less likely to attain desired outcomes. Fewer instances of covariate imbalance may have favored group care youths. In the discharge outcomes, group care youths were 6 months older on average than TFC youths. Although the relationship between urbanicity and positive outcome is unclear, it should also be noted that for follow-up outcomes, more youth in TFC were from an urban county compared to GC youth (10% vs. 3%).

Overall, the handful of post-matching imbalance generally showed that the group care youth were slightly more troubled than the TFC youth, with more previous placements and greater prevalence of domestic violence and substance abuse in the family of origin. In a few models, group care youth were about six months older on average than TFC youth and less likely to come from urban counties. Based on these group characteristics, at least, recipients of TFC may be slightly less likely to have adverse outcomes. Results from the models where covariate imbalance occurred should be interpreted with these considerations in mind.

Findings comparing outcome achievement rates are presented by outcome and matching method in Table 2. For each matching strategy, the sample size of the matched group is reported. Next, the proportion of each group that experienced the specified outcome is reported. Difference scores that are positive suggest that a higher proportion of TFC youth achieved the outcome, compared to GC youth. For negative difference scores, TFC youth were less likely to experience the outcome specified. The confidence intervals are derived from bootstrapped standard errors.

None of the findings supported the hypothesis that TFC youth would be more likely to experience positive outcomes. In fact, in two-tailed testing, TFC youth were less likely to be favorably discharged, less likely to return home, and more likely to experience subsequent formal placements. For two of the follow-up outcomes (legal involvement and homelike setting), the differences in outcome achievement across the two groups were not significant.

Despite the five different matching and stratification methods used, the findings were substantively robust across matching method. Each of the matching strategies utilized a slightly different sample size. The matched samples that were larger in size incorporated more information from the overall sample. However, findings can only be generalized to the sample included in the matched comparisons.

Considering how the findings from this study differed from previous work, post-hoc sensitivity analyses were conducted to assess how the impact of unmeasured covariates may have affected results. Mhbounds was conducted for each outcome where significant differences were found (Becker & Caliendo, 2007). Due to an assumption of independence, mhbounds estimates are biased for matching strategies that allow replacement; hence, only the results from nearest neighbor without replacement (NNWR) were included in sensitivity analyses. In this study, the NNWR matching method generally estimated smaller differences between the two groups and had smaller sample sizes, thus the results from the post-hoc analysis are likely to be more sensitive than the other matching methods.

The unobserved selection bias that would substantively change the results in measuring differences for rates of returning home or favorable discharge were rather low (1.2 and 1.7, respectively). In other words, if there were one or more unmeasured variables that increased the odds of receiving TFC by a factor of 1.2 or 1.7, differences in outcome attainment would no longer be significant. Hence, at relatively low levels of hidden bias, the confidence intervals of the treatment effect would include zero. The likelihood of experiencing a subsequent formal placement was least sensitive to hidden bias. For this outcome, one or more variables would need to more than triple the odds of receiving TFC. Based on these post-hoc analyses, it is possible that findings for some of the outcomes may alter with the inclusion of other relevant background variables.

Discussion

This study compared outcomes of youth who received treatment foster care with youth in group care. Unlike previous studies, both programs compared in this study were administered by the same organization, Girls and Boys Town, and followed the same theoretical framework, the Teaching-Family model. It was hypothesized that youth in TFC would have better outcomes; however, no support was generated for the study hypotheses. Group care youth were more likely to be favorably discharged, more likely to return home, and less likely to experience a subsequent formal placement. No differences were found in subsequent legal involvement or the likelihood of living in a homelike setting six months after discharge. These perplexing results called for additional investigation. Several possible explanations may shed light on these results.

First, it is quite possible that the two groups were different on one or more background characteristics that were not measured in the study and that affected the outcome. For example, having a sibling who is also in care would likely decrease the likelihood of returning home and increase the probability of a subsequent placement (ideally, one in which the siblings were placed together). Additional unmeasured variables, including length of stay in previous placement and placement goals may also have differed significantly in these two samples. In this sample, the TFC youth more clearly resembled a child welfare population while only a small portion of the group care youth were child welfare wards. Issues of delinquency were much more common in the group care population. Anti-social behaviors related to delinquency may be much more amenable to behaviorally-oriented treatment interventions than systemic family problems common among families receiving child welfare services. These findings provide a reminder that while delinquency-related issues may be effectively addressed in out-of-home treatment, placement is not sufficient to resolve the challenges faced by multi-problem families.

Differences in favorable discharge status across the two groups could be due to dissimilar evaluation criteria of placement success for TFC and group care. There may be pressures for group care staff to “game” this outcome measure and liberally award favorable discharge status. However, subsequent analyses did show a relationship between discharge status and follow-up outcomes, so this measure is not likely to be entirely meaningless.

Finally, findings from this study shine light on the possibility that there may be some youth who benefit from a structured group care program. While these findings should not be used to argue that group care is more effective than treatment foster care, they do call for further investigation on whether well-developed residential programs can be more effective for some youth than a less restrictive community setting. In many ways, the group care program at Girls and Boys Town is not the typical residential programs used as comparisons in earlier studies. Unique features of the Teaching-Family model may have led to improved outcomes for youth.

Limitations

Because the results in this study are dramatically different from other studies of group care and TFC, caution should be taken in interpreting these findings. Limits of the methodology may be one important consideration. The matching efforts may not have adequately accounted for all the possible differences between the two groups, as propensity score matching can only match on observed variables. A few covariates remained imbalanced even after matching with some algorithms. Another limitation of the selection bias correction strategy used here is that each of the covariates are given the same weight in calculation of the propensity score; however, it is likely that some covariates are more predictive of treatment selection and outcome than other covariates. There is currently no standard way to account for this issue in the methodology.

Common to studies that utilize administrative data, there are some measurement limitations in this project. Standardized measures of youth behavior problems and mental health need were not available in this study. The available data concerning the history of youth behavior and family problems are based on self-reports of the youth, parent, or placement worker. These individuals may have a vested interest in minimizing or overstating actual need to promote or prevent placement. Many of the covariates used in the propensity score matching analyses are prone to these biases. Further, the outcomes measured are typical indicators found in administrative data, but lack the richness of more clinically relevant measures of functioning. Several of the selected outcomes are also impacted by systems issues that reflect more than just the youth's improvement during placement.

This sample of youth in residential care at Girls and Boys Town may not represent youth in residential programs nationwide. Average length of stay in the residential program at GBT is longer than many other programs, and youth are less likely to be in public custody compared to study populations in other group care research. In addition, the treatment services delivered through Girls and Boys Town are not likely to be typical of all TFC or group care programs. Girls and Boys Town requires a high level of training for all staff and foster parents and provides significant oversight and consultation on the application of the Teaching-Family model. Other programs that utilize shift staff or a different therapeutic approach may not experience the same outcomes as this study.

Although the treatment received by all youth in each treatment type was considered homogenous because the intervention is standardized, differences in the way the treatment model is utilized may exist. Services received by youth in treatment foster care may differ by program site or foster home. Despite training and ongoing supervision to help families adhere to the program model, it is possible that individual staff or foster parents may emphasize some components of the treatment model more than others. Characteristics of staff and foster parents, including personality, experience, age, race, gender, and sophistication in using the model, likely vary. A youth's relationship with a caregiver as well as the caregiver's capacity to care for challenging youth may also impact youth outcomes. However, these constructs were not considered in this analysis.

Implications

If this study had found more prevalent negative outcomes for group care youth, it would be possible to conclude that even one of the most resource-rich, state-of-the-art family-style group care models was inferior to treatment foster care. These results would have provided additional ammunition to individuals calling for the dismantling of group care practice. However, the findings suggest that an effective means for providing group care services is available.

One of the theoretical reasons that TFC was expected to produce better outcomes than group-based treatment is that group care exposes youth to other deviant peers. Much debate has centered on the iatrogenic effects of peer contagion (see Dishion, McCord & Poulin, 1999 or Dodge, Dishion & Landsford, 2006). However, in this study, group care youths were not less likely to experience positive outcomes than the youth placed individually in specialized foster homes. Aspects of the Teaching-Family model, including high levels of adult supervision and ample pro-social activities, may have minimized any negative peer effects. Future research examining evidence for peer contagion in structured, family-style group care programs should be conducted.

In considering the possibility that effective group care is attainable, the focus of research efforts should shift to identifying the critical ingredients in creating high quality group care programs. The findings from this study suggest that elements of the Teaching-Family model may offer clues to what works in group care settings. Exploratory work to identify quality indicators of group care programs could lead to the development of measures to assess the quality and effectiveness of group care settings.

Placement decision-making for youth in public systems embraces the policy of least restrictive placement. This study suggests that for some youth, placement into group care may be more effective than receiving TFC. More research is needed to understand which youth benefit from a more restrictive placement and which youth would be better served in a less restrictive setting. Currently, placement decision-making is left to child welfare workers and courts and little is known about how these decisions are made. Studies finding geographic variation and racial disparity in group care placements suggest that these decisions are not based on individual youth needs, but rather system or placement worker characteristics (McMillen et al., 2005). A better understanding of what works for which youth will likely increase the efficient allocation of placement resources.

Acknowledgements

The preparation of this article was supported by the Center for Mental Health Services Research at the Brown School of Social Work, Washington University in St. Louis, and through award 5P30-MH-068579 from the National Institute of Mental Health. The authors thank Curtis McMillen, Ph.D., Rick Barth, Ph.D., and Elizabeth Caplick Weigensburg for their input on this study. Comments from the anonymous reviewers also greatly strengthened this manuscript.

References

- Barth, RP. Institutions vs. foster homes: The empirical base for a century of action. Chapel Hill, NC: UNC, School of Social Work, Jordan Institute for Families; 2002.
- Barth RP. Foster home care is more cost-effective than shelter care: Serious questions continue to be raised about the utility of group care in child welfare services. *Child Abuse & Neglect* 2005;29:623–625. [PubMed: 15979705]
- Bates BC, English DJ, Kouidou-Giles S. Residential treatment and its alternatives: A review of the literature. *Child & Youth Care Forum* 1997;26(1):7–51.
- Becker SO, Caliendo M. Sensitivity analysis for average treatment effects. *The Stata Journal* 2007;7(1): 71–83.
- Berrick JD, Courtney M, Barth RP. Specialized foster care and group home care: Similarities and differences in the characteristics of children in care. *Children and Youth Services Review* 1993;15:453–473.
- Breland-Noble AM, Elbogen EB, Farmer EMZ, Dubs MS, Wagner HR, Burns BJ. Use of psychotropic medications by youths in therapeutic foster care and group homes. *Psychiatric Services* 2004;55(6): 706–707. [PubMed: 15175471]
- Breland-Noble AM, Farmer EMZ, Dubs MS, Potter E, Burns BJ. Mental health and other service use by youth in therapeutic foster care and group homes. *Journal of Child and Family Studies* 2005;14(2): 167–180.

- Braitman LE, Rosenbaum PR. Rare outcomes, common treatments: Analytic strategies using propensity scores. *Annals of Internal Medicine* 2002;137(8):693–695. [PubMed: 12379071]
- Chamberlain, P.; Mihalic, SF. Multidimensional Treatment Foster Care: Blueprints for Violence Prevention, Book Eight. In: Elliott, DS., editor. *Blueprints for Violence Prevention Series*. Boulder, CO: Center for the Study and Prevention of Violence, Institute of Behavioral Science, University of Colorado; 1998.
- Chamberlain P, Ray J, Moore KJ. Characteristics of residential care for adolescent offenders: A comparison of assumptions and practices in two models. *Journal of Child and Family Studies* 1996;5(3):285–297.
- Chamberlain P, Reid JB. Using a specialized foster care community treatment model for children and adolescents leaving the state mental hospital. *Journal of Community Psychology* 1991;19:266–276.
- Chamberlain P, Reid JB. Comparison of two community alternatives to incarceration of chronic juvenile offenders. *Journal of Consulting and Clinical Psychology* 1998;66(4):624–633. [PubMed: 9735578]
- Child Welfare League of America. *CWLA Standards of Excellence for Residential Services*. Washington, DC: Child Welfare League of America; 2004.
- Coen, AS.; Libby, AM.; Price, DA.; Silverman, K. *Inside the black box: A study of the residential treatment center program of Colorado*. Denver, CO: Division of Child Welfare Services, Colorado Department of Human Services; 2003.
- Collins ME, Schwartz IM, Epstein I. Risk factors for adult imprisonment in a sample of youth released from residential child care. *Children and Youth Services Review* 2001;23(3):203–226.
- Curtis PA, Alexander G, Lunghofer LA. A literature review comparing the outcomes of residential group care and therapeutic foster care. *Child and Adolescent Social Work Journal* 2001;18(5):377–392.
- D’Agostino RB. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statistics in Medicine* 1998;17:2265–2281. [PubMed: 9802183]
- Dahl E, Lorentzen T. What works for whom? An analysis of active labour market programmes in Norway. *International Journal of Social Welfare* 2005;14:86–98.
- DeSena AD, Murphy RA, Douglas-Palumberi H, Blau G, Blandina Horwitz SM, Kaufman J. SAFE Homes: Is it worth the cost? An evaluation of a group home permanency planning program for children who first enter out-of-home care. *Child Abuse & Neglect* 2005;29:627–643. [PubMed: 15979706]
- Dishion TJ, McCord J, Poulin F. When interventions harm. *American Psychologist* 1999;54(9):755–764. [PubMed: 10510665]
- Dodge, KA.; Dishion, TJ.; Landsford, JE., editors. *Deviant peer influences in programs for youth*. New York: Guilford Press; 2006.
- Dodge, KA.; Sherrill, MR. Deviant peer group effects in youth mental health interventions. In: Dodge, KA.; Dishion, TJ.; Landsford, JE., editors. *Deviant peer influences in programs for youth*. New York: Guilford Press; 2006. p. 97-120.
- Drais-Parrillo, AA., editor. *The Odyssey Project: A descriptive and prospective study of children and youth in residential group care and therapeutic foster care*. Washington, DC: Child Welfare League of America; 2004.
- Eddy JM, Chamberlain P. Family management and deviant peer association as mediators of the impact of treatment condition on youth antisocial behavior. *Journal of Consulting and Clinical Psychology* 2000;68(5):857–863. [PubMed: 11068971]
- El-Bassel N, Gilbert L, Wu E, Go H, Hill J. HIV and intimate partner violence among methadone-maintained women in New York City. *Social Science & Medicine* 2005;61:171–183. [PubMed: 15847970]
- Essama-Nssah, B. Propensity score matching and policy impact analysis. Paper presented at The World Bank, Poverty Reduction Group Meeting; Washington, DC. 2006.
- Fisher PA, Chamberlain P. Multidimensional treatment foster care: A program for intensive parenting, family support, and skill building. *Journal of Emotional & Behavioral Disorders* 2000;8(3):155–165.
- Fixsen, D.; Blase, K. *Publications regarding the Teaching-Family Model*. Louis de la Parte Florida Mental Health Institute; University of South Florida; 2002.
- Fraser MW. Intervention research in social work: Recent advances and continuing challenges. *Research on Social Work Practice* 2004;14(3):210–222.

- French KM, Cameron G. Treatment of choice or a last resort? A review of residential mental health placements for children and youth. *Child & Youth Care Forum* 2002;31(5):307–339.
- Friman, PC. Behavioral, family-style residential care for troubled out-of-home adolescents: Recent findings. In: Austin, J.; Carr, JE., editors. *Handbook of Applied Behavioral Analysis*. Reno, NV: Context Press; 2000. p. 187-207.
- Friman PC, Jones M, Smith GL, Daly DL, Larzelere RE. Decreasing disruptive behavior by adolescent boys in residential care by increasing their positive to negative interactional ratios. *Behavior Modification* 1997;21:470–486. [PubMed: 9337602]
- Friman PC, Toner C, Soper S, Sinclair J, Shanahan D. Maintaining placement for troubled and disruptive adolescents in voluntary residential care: The role of reduced youth-to-staff ratio. *Journal of Child and Family Studies* 1996;5(3):337–347.
- Gorske TT, Srebalus DJ, Walls RT. Adolescents in residential centers: characteristics and treatment outcome. *Children and Youth Services Review* 2003;25(4):317–326.
- Guo S, Barth RP, Gibbons C. Propensity score matching strategies for evaluating substance abuse services for child welfare clients. *Children and Youth Services Review* 2005;28(4):357–383.
- Guo, S.; Rose, RA.; Bowen, GL. Quasi-Experimental Strategies When Randomization Fails: Propensity Score Matching and Sensitivity Analysis. 2006, January; Paper presented at the Society for Social Work and Research Tenth Annual Conference; San Antonio, TX.
- Hair HJ. Outcomes for children and adolescents after residential treatment: A review of research from 1993 to 2003. *Journal of Child and Family Studies* 2005;14(4):551–575.
- Handwerk ML, Field CE, Friman PC. The iatrogenic effects of group intervention for antisocial youth: Premature extrapolations? *Journal of Behavioral Education* 2001;10(4):223–238.
- Iezzoni, L. Risk adjustment for measuring health care outcomes. Chicago: Health Administration Press; 2003.
- Larzelere RE, Daly DL, Davis JL, Chmelka MB, Handwerk ML. Outcome evaluation of Girls and Boys Town's family home program. *Education and Treatment of Children* 2004;27(2):130–149.
- Leow C, Marcus S, Zanutto E, Boruch R. Effects of advanced course-taking on math and science achievement: Addressing selection bias using propensity scores. *American Journal of Evaluation* 2004;25(4):461–478.
- Leuven, E.; Sianesi, B. Psmatch2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. 2003. <http://ideas.repec.org/c/boc/bocode/s432001.html>, version 3.0.0
- Leve LD, Chamberlain P. Association with delinquent peers: Intervention effects for youth in the juvenile justice system. *Journal of Abnormal Child Psychology* 2005;33(3):339–347. [PubMed: 15957561]
- Leve LD, Chamberlain P, Reid JB. Intervention outcomes for girls referred from juvenile justice: Effects on delinquency. *Journal of Consulting and Clinical Psychology* 2005;73(6):1181–1185. [PubMed: 16392991]
- Lipsey, MW.; Wilson, DB. Effective intervention for serious juvenile offenders: A synthesis of research. In: Loeber, R.; Farrington, DP., editors. *Serious & violent juvenile offenders: Risk factors and successful interventions*. Thousand Oaks, CA: Sage; 1998. p. 313-345.
- McMillen JC, Zima BT, Scott LD Jr, Auslander WF, Munson MR, Ollie MT, Spitznagel EL. Prevalence of psychiatric disorders among older youths in the foster care system. *Journal of the American Academy of Child and Adolescent Psychiatry* 2005;44(1):88–96. [PubMed: 15608548]
- Mitra N, Indurkha A. A propensity score approach to estimating the cost-effectiveness of medical therapies from observational data. *Health Economics* 2005;14:805–815. [PubMed: 15791679]
- Osgood, DW.; Bridell, LO. Peer effects in Juvenile Justice. In: Dodge, KA.; Dishion, TJ.; Landsford, JE., editors. *Deviant peer influences in programs for youth*. New York: Guilford Press; 2006. p. 141-161.
- Pecora, PJ.; Whittaker, JK.; Maluccio, AN.; Barth, RP.; Plotnick, RD. *The child welfare challenge: Policy, practice and research*. New York: Aldine de Gruyter; 1992.
- Piotrkowski CS, Baker AJL. Predicting discharge dispositions among adolescents in residential treatment. *Residential Treatment for Children & Youth* 2004;21(3):69–86.
- Rosenbaum PR, Rubin DR. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;70(1):41–55.

- Rosenbaum PR, Rubin DR. The bias due to incomplete matching. *Biometrics* 1986;41(1):103–116. [PubMed: 4005368]
- Ryan JP. Dependent youth in juvenile justice: Do positive peer culture programs work for victims of child maltreatment? *Research on Social Work Practice* 2006;16(5):511–519.
- Seeger JD, Williams PL, Walker AM. An application of propensity score matching using claims data. *Pharmacoepidemiology and Drug Safety* 2005;14:465–476. [PubMed: 15651087]
- Sosin MR. Outcomes and sample selection: The case of a homelessness and substance abuse intervention. *British Journal of Mathematical and Statistical Psychology* 2002;55:63–91. [PubMed: 12034012]
- Stroul, BA.; Friedman, RM. A system of care for severely emotionally disturbed children & youth. Washington, DC: CASSP Technical Assistance Center, Georgetown University Child Development Center; 1986.
- Sunseri PA. The prediction of unplanned discharge from residential treatment. *Child & Youth Care Forum* 2001;30(5):283–303.
- Teare JF, Larzelere RE, Smith GL, Becker CY, Castrianno LM, Peterson RW. Placement stability following short-term residential care. *Journal of Child and Family Studies* 1999;8(1):59–69.
- Thompson RW, Smith GL, Osgood DW, Dowd TP, Friman PC, Daly DL. Residential care: A study of short and long-term educational effects. *Children and Youth Services Review* 1996;18:221–242.
- U.S. Department of Health and Human Services. *Mental Health: A Report of the Surgeon General*. Rockville, MD: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services, National Institutes of Health, National Institute of Mental Health; 1999.
- U.S. Department of Health and Human Services. *National Survey of Child and Adolescent Well-Being: One Year in Foster Care Report*. Washington, D.C: Administration for Children, Youth and Families; 2001a.
- U.S. Department of Health and Human Services. *Youth Violence: A Report of the Surgeon General*. Rockville, MD: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services, National Institutes of Health, National Institute of Mental Health; 2001b.
- U.S. Department of Health and Human Services. *The AFCARS Report, Appendix D*. 2003 [Retrieved October 5, 2003]. from <http://www.acf.hhs.gov/programs/cb>
- Vorrath, HH.; Brendtro, LK. *Positive peer culture*. 2. New York: Aldine Publishing; 1985.
- Wells K. Placement of emotionally disturbed children in residential treatment: A review of placement criteria. *American Journal of Orthopsychiatry* 1991;61(3):339–347. [PubMed: 1951641]
- Whittaker JK. Reinventing residential childcare: An agenda for research and practice. *Residential Treatment for Children & Youth* 2000;17(3):13–30.
- Winship C, Morgan SL. The estimation of causal effects from observational data. *Annual Review of Sociology* 1999;25:659–706.
- Wolf MM, Phillips EL, Fixsen DL, Braukmann CJ, Kirigin KA, Willner AG, Schumaker J. *Achievement Place: The Teaching-Family model*. *Child Care Quarterly* 1976;5(2):92–103.
- Youngbauer, JG. *The Teaching-Family model and treatment durability: Assessing generalization using survival analysis techniques*, Unpublished doctoral dissertation. University of Kansas; Lawrence: 1997.

Table 1
Pre and Post-matching covariate differences for Youth in Treatment Foster Care and Group Care

	Treatment Foster Care	Group Care	Pre-match t-value/ χ^2	Post-match
Total Sample	112	716		
Male	43 (38%)	442 (62%)	21.91***	ns
White	57 (51%)	429 (60%)	3.31***	ns
Neglect history	76 (67%)	291 (41%)	28.9***	ns
Physical abuse history	36 (32%)	198 (28%)	.945**	ns
Sexual abuse history	33 (29%)	121 (17%)	10.05***	ns
Legal delinquent status	22 (20%)	626 (88%)	263.31***	ns
Formal (Child welfare, Juvenile Justice, Mental Health)	91 (81%)	304 (43%)	58.22***	ns
Previous detention placement	2 (2%)	106 (15%)	14.5***	ns
Youth Behavior Problem Index score	2.1 (1.8)	3.2 (2.1)	5.04***	ns
County Population			10.54**	ns ¹
Urban	9 (8%)	56 (8%)		
Small city	34 (31%)	333 (47%)		
Non-metropolitan	68 (61%)	325 (45%)		
Parent mental health issues	51 (45%)	143 (20%)	35.17***	ns
Parent legal involvement	39 (35%)	201 (28%)	2.18**	ns
Incarcerated parent	29 (26%)	113 (16%)	6.93**	ns
Parent with substance issues	21 (19%)	150 (21%)	.29*	ns ¹
Domestic violence	46 (41%)	223 (31%)	4.31*	ns ¹
Mean Age at Intake in Years(SD)	12.9 (2.5)	14.9 (1.7)	11.22***	ns ²
Length of Stay in Months (SD)	14.6 (14.9)	20.3(14.8)	4.036***	ns
Number of Prior Placements (SD)	3.4 (3.5)	2.0 (2.6)	-4.67***	ns ²

* p<.05,

** p<.01,

*** p<.001

¹ significant differences found with one of the four matching methods

² significant differences found with two of the four matching methods

Table 2
Estimated effect of TFC on outcomes using different matching methods

	N TFC	N GC	TFC	GC	Diff.	95% CI
Favorable discharge						
NN no replacement	112	112	.60	.81	-.21*	-.36, -.14
NN with replacement	112	45	.60	.93	-.33***	-.41, -.27
NN 1:6	112	136	.60	.91	-.31***	-.44, -.24
Kernel	112	716	.60	.89	-.29***	-.39, -.22
Return home						
NN no replacement	110	110	.40	.56	-.16***	-.30, -.04
NN with replacement	110	48	.40	.65	-.25**	-.48, -.19
NN 1:6	110	169	.40	.69	-.25***	-.41, -.12
Kernel	110	714	.40	.69	-.29**	-.46, -.14
Follow-up Legal involvement						
NN no replacement	76	76	.12	.13	-.01	-.11, .15
NN with replacement	76	41	.12	.08	.04	-.04, .14
NN 1:6	76	106	.12	.10	.02	-.14, .09
Kernel	76	542	.12	.11	.01	-.11, .09
Subsequent formal placements						
NN no replacement	85	85	.70	.27	.43***	.23, .53
NN with replacement	85	44	.69	.15	.54***	.42, .71
NN 1:6	85	111	.70	.19	.51***	.31, .65
Kernel	85	417	.70	.19	.51***	.23, .62
Home-like setting at follow-up						
NN no replacement	85	85	.71	.62	.08	-.09, .20
NN with replacement	85	44	.71	.65	.06	-.17, .27
NN 1:6	85	118	.71	.62	.09	-.07, .34
Kernel	85	567	.71	.61	.09	-.12, .23

* p<.05,

** p<.01,

*** p<.001