Journal of Consulting and Clinical Psychology
2010, Vol. 78, No. 2, 144–155


0022-006X/10/$12.00 DOI: 10.1037/a0018544

# Youth Psychotherapy Change Trajectories and Outcomes in Usual Care: Community Mental Health Versus Managed Care Settings

Jared S. Warren, Philip L. Nelson, Sasha A. Mondragon, Scott A. Baldwin, and Gary M. Burlingame
Brigham Young University

***Objective:*** The authors compared symptom change trajectories and treatment outcome categories in children and adolescents receiving routine outpatient mental health services in a public community mental health system and a private managed care organization. ***Method:*** Archival longitudinal outcome data from parents completing the Youth Outcome Questionnaire (Y-OQ) were retrieved for children and adolescents (4–17 years old) served in a community mental health system ($n$ = 936, mean age = 12 years, 40% girls or young women, 28% from families of color) and a managed care organization ($n$ = 3,075, mean age = 13 years, 45% girls or young women, race and ethnicity not reported). The authors analyzed Y-OQ data using multilevel modeling and partial proportional odds modeling to test for differences in change trajectories and final outcomes across the 2 service settings. ***Results:*** Although initial symptom level was comparable across the 2 settings, the rate of change was significantly steeper for cases in the managed care setting. In addition, 24% of cases in the community mental health setting demonstrated a significant increase in symptoms over the course of treatment, compared with 14% of cases in the managed care setting. ***Conclusions:*** These results emphasize the need for increased attention to negative outcomes in routine mental health services and provide a stronger foundation for identifying youth cases at risk for treatment failure. In addition, given the overall differences observed across treatment settings for average rate of change and deterioration rates, results suggest that setting-specific model heuristics should be used for identifying cases at risk for negative outcomes.

*Keywords:* youths, psychotherapy, change trajectories, outcome, usual care

The study of treatment outcomes and change processes in traditional community practice settings is one of the most urgent needs in children's mental health services (Burns, Hoagwood, & Mrazek, 1999; Weisz & Gray, 2008; Weisz, Jensen, & McLeod, 2005). The small body of outcome studies in community-based usual care settings has yielded sobering results, with overall mean effect sizes near zero (Weiss, Catron, Harris, & Phung, 1999; Weisz, 2004; Weisz, Donenberg, Han, & Weiss, 1995), yet millions of youth are served each year in these systems of care (National Advisory Mental Health Council, 2001; Ringel & Sturm, 2001). Furthermore, although the broader research base on efficacious treatments for children and adolescents is impressive, significant concerns exist regarding the applicability and generalizability of these studies to usual clinical care (Garland, Hurlburt, & Hawley, 2006; Weisz, 2004; Weisz et al., 2005). Citing the many differences in patients, therapists, and treatment conditions typically observed between research-based services and usual clinical care, Weisz and colleagues (2005) indicated that the most valid answers to questions about treatment outcome and change processes are more likely to come from research in "real-world" service settings than from controlled lab studies. Consequently, there is a strong need for careful examination of treatment outcomes and change processes in usual clinical care.

## Change Trajectories in Psychotherapy

A central issue in understanding youth psychotherapy outcomes and therapeutic processes is the study of change over time. That is, in addition to knowing that change occurs in response to treatment, it is important to understand *patterns* of change over the course of treatment. Examining the course of change in symptom level—the *shape* of change over time—yields valuable information for clinical research, practice, and theory.

One of the most promising clinical applications of research in this area is the use of empirically derived change trajectories as a foundation for evaluating individual patient progress. Using the "patient-focused research" paradigm (Howard, Moras, Brill, Martinovich, & Lutz, 1996), several researchers have devised methods to track the progress of individual patients and to alert clinicians when a patient may be at risk for treatment failure or premature dropout (Finch, Lambert, & Schaalje, 2001; Howard et al., 1996; Lambert, Hansen, & Finch, 2001). In the most accurate of these methods, empirically derived warning systems based on expected

Jared S. Warren, Sasha A. Mondragon, Scott A. Baldwin, and Gary M. Burlingame, Department of Psychology, Brigham Young University; Philip L. Nelson, Department of Counseling Psychology and Special Education, Brigham Young University.

Gary M. Burlingame is a co-developer of the Youth Outcome Questionnaire (Y-OQ) used in this study; he is also a principal in OQ Measures, distributor of the Y-OQ.

We express sincere appreciation to Catherine Carter, Dave Justice, and Mark Latkowski for their assistance in this research, particularly in providing administrative support and database management. This work would not be possible without their valued contributions.

Correspondence concerning this article should be addressed to Jared S. Warren, Department of Psychology, Brigham Young University, 291 TLRB, Provo, UT 84602. E-mail: jared_warren@byu.edu

change trajectories are used, and the characteristics or the initial level of symptom severity of individual patients often are taken into account (Lambert et al., 2002). For example, Lambert and colleagues (Finch et al., 2001; Lambert et al., 2001) developed a system in which individual patient progress can be compared on a session-by-session basis with expected change trajectories modeled from outcome data of hundreds of patients with similar initial symptom levels. If the patient's symptom level at a given session during treatment deviates considerably from the average (expected) change trajectory, an "alert" is provided to the clinician, prompting more focused assessment and re-evaluation of the approach to treatment for that patient. Clinical trials of these methods with adult patients in a university-based counseling center have yielded impressive results in reducing rates of treatment failure and minimizing premature dropout from treatment (Harmon et al., 2007; Hawkins, Lambert, Vermeersch, Slade, & Tuttle, 2004; Lambert et al., 2001). However, because expected change trajectories are likely to vary across different treatment settings, modalities, and presenting problems, such warning systems may need to be tailored to match the conditions of the treatment setting.

Research on expected change trajectories may also provide insurance companies and other third-party payers with better information regarding the number of sessions typically required for improvement and recovery in patients with various levels and forms of psychopathology. In addition, identifying factors that account for individual variation in patient change trajectories could prompt third-party payers to be more flexible in their decisions about the number of sessions allowed or the level of services covered for a particular patient. Consequently, studying typical patterns of change in psychotherapy—and deviations from such trajectories—holds considerable promise for improving routine clinical services.

Examining the course of symptom progression during treatment also has important implications for testing and refining theories of change, as different interventions are often hypothesized to produce different patterns of change over time (e.g., psychoeducational vs. exposure-based interventions; Laurenceau, Hayes, & Feldman, 2007). Very little attention has been devoted to how well actual patient change over the course of treatment resembles change predicted by theory, and the great majority of existing outcome studies are unable to address such questions because the typical pre/post/follow-up design prohibits examination of actual patterns of change during treatment (Laurenceau et al., 2007). Accordingly, the study of change trajectories in mental health treatment may be particularly well suited to respond to calls for increased integration of theory and research in mental health services (Kazdin, 2000; Silverman, Kurtines, & Hoagwood, 2004).

## Youth Outcomes in Usual Care

As previously noted, research summarizing the average impact of child psychotherapy in usual care has yielded effect sizes near zero (Weiss et al., 1999; Weisz, 2004). Although these important findings emphasize the need for improved dissemination and deployment of efficacious interventions for youths in "real-world" treatment settings, they can potentially obscure the fact that considerable variation exists in individual response to treatment. Simply stated, some youths served in usual care settings demonstrate significant improvement in symptoms, some show no reliable change, and a sizable minority leave treatment significantly worse off than when they began. The latter group—those experiencing *deterioration* or *treatment failure*—is of particular interest and concern given the service provider's ethical credo to "do no harm" (Lilienfeld, 2007). In adult psychotherapy outcome research, approximately 5–10% of patients leave treatment worse off than when they began (Lambert & Bergin, 1994; Mohr, 1995); however, deterioration rates may exceed 20% in youths served in traditional community mental health settings (Warren, Nelson, & Burlingame, 2009).

Evaluating youth outcomes with reliable change criteria (Jacobson & Truax, 1991) provides a useful supplement to traditional methods of examining average group response to treatment. Examining response to treatment by outcome class (e.g., patients who demonstrate reliable improvement, no reliable change, or deterioration) also permits more focused attention on specific categories of outcome such as deterioration (i.e., treatment failure). For example, data on the percentage of patients experiencing deterioration are often used in conjunction with empirical change trajectories as an important component in systems for identifying cases at risk for treatment failure (Lambert et al., 2002; Warren et al., 2009). As was noted previously with regard to the use of expected change trajectories in such warning systems, if treatment failure rates vary significantly across different treatment settings, this warning system component may need to be adjusted to match the specific setting in which the system is applied.

## Community Mental Health Versus Managed Care Settings

Relatively little research has been conducted on child and adolescent mental health outcomes and change processes in usual care settings. This deficiency is magnified by the fact that youths may receive mental health services through different types of providers and systems, depending on services available in the community and the family's financial resources, among other factors. For example, two of the most common settings in which youths receive mental health services in the United States are public community mental health systems and private managed care organizations. These two major types of service systems share many common features and may serve the same population of families in an area; however, a number of distinctions are typically observed. For example, community mental health centers in the United States are often supported at least in part by government assistance programs such as Medicaid, allowing them to serve more youths from lower income families. In contrast, private managed health care organizations are more likely to serve youths from families who can afford private health insurance or who have employer-provided health benefits. As a result, community mental health systems often serve families with more severe constellations of stressors and negative circumstances that often accompany financial disadvantage. In addition, community mental health treatment more often emphasizes a "system-of-care" approach in which supplemental services such as case management, group treatment, and residential treatment are more closely integrated (Center for Mental Health Services, 1999). In contrast, a patient served through a private managed care system is more likely to be treated by a single mental health professional (who is a member of the system's network panel of providers) and is more likely to receive individ-

ual therapy only. Furthermore, given the increased emphasis in managed care on providing quality services while minimizing costs, the average duration of treatment in managed care is often shorter than in traditional community mental health settings. This notion is reinforced by a common criticism of the traditional managed care model as being ill suited for treating persons with chronic mental health conditions (Anderson, 2007).

The differences between public community mental health systems and private managed care organizations may have important implications for child and adolescent treatment outcomes and change processes. For example, it is reasonable to assume that if community mental health systems typically serve more lower income families with more pervasive life stressors and disadvantages, this difference could be reflected in the average patient symptom level at the beginning of treatment, as well as the average rate of change and final outcome of treatment. In addition, differences in the average change trajectories and deterioration rates across these two settings would suggest that different model heuristics should be used to identify cases at risk for treatment failure in these settings. However, no study to date has examined potential differences in initial symptom level, pattern and rate of change, and final outcome across these two major mental health settings.

## Study Aims

The purpose of this study was to compare youths' initial symptom levels, patterns of change, and final outcome in a large community mental health system and a managed care organization that serve families in the same geographic area. In addition, the results of this study were subsequently used as the foundation for a companion study on the accuracy of a system for identifying cases at risk for treatment failure (Warren, Nelson, Burlingame, & Mondragon, 2010). Given the presumption that youths from lower income families with more complex constellations of stressors and fewer resources are served more frequently in community mental health systems than in private managed care organizations, our primary hypotheses were as follows: (a) initial symptom levels would be higher in the community mental health setting, (b) the average rate of change would be steeper in the managed care setting, and (c) negative outcomes (i.e., no reliable change or deterioration) would be more frequent in the community mental health setting.

## Method

### Participants and Procedures

We analyzed data selected from the archives (1997–2008) of a public community mental health system and a large private managed care organization, both located in the intermountain western United States. The study was conducted in compliance with appropriate institutional review boards.

The community mental health system serves an area of approximately 1.5 million residents, with clientele typically of average to low socioeconomic status (SES). The professional backgrounds of the 175 clinicians from this setting who provided psychotherapy services to patients in this study were as follows: 73% social workers, 9% psychologists, 6% licensed professional counselors, 3% psychiatrists (included only if providing psychotherapy), 2% marriage and family therapists, and 7% other/unknown. Services included individual and family psychotherapy, psychosocial skill-building groups, and medication management visits. The type of treatment provided by therapists in this setting is best described as "eclectic," similar to the "clinically derived treatment" in usual clinical care described by Weisz (2004, p. 12; i.e., common elements included talking or playing with the child, talking with parents about concerns, listening reflectively, showing empathy, responding to the issues the child or parent brings to therapy, and so on). When asked about specific techniques and orientations used, therapists generally reported that they employed family therapy and cognitive strategies more frequently than psychodynamic or behavioral techniques; however, they did report using a fairly broad range of therapeutic approaches in the community mental health setting. Therapists in this setting were expected to accrue 20–25 billable patient hours per week and reportedly maintained a typical open caseload of 90–110 patients.

The managed care organization primarily serves families of average to high SES with private health insurance, covering approximately 1 million individuals. Families not insured by the managed care organization are typically referred elsewhere for mental health treatment (such as the community mental health system in this study). Mental health treatment provided by the managed care organization was described by clinical supervisors as being multidisciplinary and eclectic, with short-term cognitive–behavioral interventions being most frequently emphasized and encouraged by supervisors. Individual psychotherapy, family therapy, group therapy, and medication management were the most common modes of treatment provided. Parent involvement in youth treatment was reported to be strongly emphasized. The professional backgrounds of the 55 clinicians from this setting who provided services to patients in this study were as follows: 27% social workers, 18% psychiatrists, 13% psychologists, 4% licensed professional counselors, and 38% other/unknown. Therapists in the managed care setting were expected to accrue at least 25 billable patient hours per week and reportedly maintained an open caseload of approximately 100 patients.

Within each institution, outcome data were collected as part of routine clinical services. Parents or guardians completed the Youth Outcome Questionnaire (Y-OQ, Burlingame et al., 2001; Burlingame et al., 2005; Burlingame, Wells, Lambert, & Cox, 2004; described later) at check-in when presenting their children for outpatient treatment; the questionnaire typically required less than 10 min to complete. The number of cases from the original archives with at least one Y-OQ totaled 3,524 from the community setting and 4,364 from the managed care setting. We selected cases within the appropriate Y-OQ age range of 4–17 years, which was 99% of the community archive and 95% of the managed care archive. Selecting cases with at least two measurement occasions further reduced our sample to 28% from the community setting and 76% from the managed care setting. Selecting cases not having extremely long treatment episodes (i.e., below the 90th percentile of cases with at least one Y-OQ for each setting: 83 weeks or less for community cases and 25 weeks or less for managed care cases) further reduced our sample to 27% and 70%, respectively. Tables 1A and 1B present descriptive statistics for our selected sample of 936 community system cases and 3,075 managed care cases.

For the community setting, the selected sample differed significantly from the original archive, with a slightly lower mean age

Table 1A
*Descriptive Statistics for Samples Selected From Community and Managed Care Settings*

| Variable | Community ($n$ = 936) | | | Managed care ($n$ = 3,075) | | | Community and managed care comparison | |
|---|---|---|---|---|---|---|---|---|
| | *M* | *SD* | *Mdn* | *M* | *SD* | *Mdn* | *t* | *df* |
| Y-OQs per client | 2.7 | 1.2 | 2.0 | 3.5 | 2.3 | 3.0 | 14.97* | 3,009 |
| Y-OQs per month | 0.7 | 0.9 | 0.5 | 2.6 | 5.2 | 1.6 | 19.39* | 3,560 |
| Sessions per month | 2.2 | 1.7 | 1.8 | 2.6[a] | 5.2[a] | 1.6[a] | 3.80* | 4,004 |
| Sessions per Y-OQ | 4.0 | 3.3 | 3.0 | 1.0[a] | 0.0[a] | 1.0[a] | −27.80* | 935 |
| Treatment episodes | | | | | | | | |
| No. of episodes | 1.7 | 1.8 | 1.0 | 1.4 | 1.0 | 0.8 | −5.34* | 1,057 |
| Length in weeks | 25.3 | 18.9 | 20.1 | 8.7 | 5.9 | 7.9 | −26.58* | 992 |
| Length in sessions | 11.1 | 11.2 | 8.0 | 3.5[a] | 2.3[a] | 3.0[a] | −20.60* | 959 |
| Age | 12.0 | 3.5 | 12.3 | 12.9 | 3.5 | 13.5 | 6.09* | 1,554 |

*Note*. *Mdn* = median; *df* = degrees of freedom; Y-OQs = Youth Outcome Questionnaires.
[a] These numbers are estimates. Standard protocol in the managed care setting is to administer a Y-OQ at every session.
* $p < .05$.

(11.6 vs. 12.3 years, respectively), a higher first Y-OQ total score (86 vs. 77, respectively), a longer mean treatment episode (32 weeks or 15 sessions vs. 12 weeks or 4 sessions, respectively), a greater percentage of cases whose treatment included medications (70% vs. 34%), a greater percentage of cases with serious emotional disorders (93% vs. 84%), and greater percentage of Medicaid cases (56% vs. 49%). The selected community sample did not differ from the archive in percentage of girls or young women, percentage of cases who had received prior treatment, percentage of minorities, percentage of Hispanic cases (included in minority category), or percentage of cases with documented alcohol and drug use.

For the managed care setting, the selected sample differed significantly from the original archive, with a slightly higher mean age (13.0 vs. 11.3 years, respectively), a longer mean treatment episode (12 weeks or 3.0 sessions vs. 7 weeks or 2.9 sessions), and a greater percentage of cases with previous treatment (24% vs. 5%). The selected managed care sample did not differ from the archive in mean first Y-OQ score or percentage of girls or young women.

The community and managed care settings differed significantly in most domains listed in Table 1A, with the most notable differences in measurement frequency (i.e., 0.7 Y-OQs per month for the community setting vs. 2.6 Y-OQs per month for the managed care setting) and treatment episode length (i.e., 25.3 weeks for the community setting vs. 8.7 weeks for the managed care setting). We considered a treatment episode complete after 90 days of no patient contact. Community system cases received a mean of 11.1 treatment sessions, and managed care cases received a mean of 3.5 treatment sessions. In addition, the managed care sample had a higher percentage of girls and young women than the community sample, 45% vs. 40%, $\chi^2(1) = 6.64$, $p = .01$, yet roughly the same percentage of cases with previous treatment, 27% vs. 30%, $\chi^2(1) = 3.09$, $p = .08$. Table 1B presents descriptive statistics available only in the community system archive, including primary diagnoses.

Table 1B
*Descriptive Statistics Available Only for Sample Selected From Community Setting*

| Variable | % |
|---|---|
| Primary diagnoses[a] | |
| Attention-deficit/hyperactivity disorders | 22.2 |
| Mood disorders | 20.5 |
| Substance abuse/dependence | 9.7 |
| Oppositional defiant disorder | 9.2 |
| Adjustment disorders | 8.6 |
| Posttraumatic stress disorder | 7.1 |
| Abuse/neglect of child | 6.9 |
| Conduct disorders | 4.3 |
| Anxiety-related disorders | 4.0 |
| Autistic disorders | 1.6 |
| Other/unknown | 6.0 |
| Other information | |
| Hispanic | 9.1 |
| Ethnic/racial minority (includes Hispanic) | 28.1 |
| Medicaid | 60.4 |
| Alcohol and drug use | 12.8 |
| Clients taking medications | 68.8 |
| Serious emotional disturbance classification | 91.9 |

[a] 85.7% of clients had multiple diagnoses.

## Outcome Measure

The Youth Outcome Questionnaire 2.01 (Y-OQ; Burlingame et al., 2001; Burlingame et al., 2005; Burlingame, Wells, Lambert, & Cox, 2004) is a parent- or guardian-completed questionnaire for youths ages 4–17, requiring 8–10 min for completion. Its 64 items are rated on a 5-point Likert-type scale, and summative scoring is used to produce a total score for overall psychosocial distress. Total scores may range from −16 to 240 (eight reverse-scored items measuring adaptive behaviors can produce negative scores), with higher scores indicating greater psychosocial distress. Scores at or above the established clinical cutoff score of 46 are considered in the clinical range for level of distress. Burlingame et al. (2005) reported mean total scores for outpatient community mental health and managed care samples ranging from 68 to 76, with a

nonclinical community sample mean of 21. Although only the Y-OQ total score was used in the analyses, the Y-OQ's items also form six subscales corresponding to behavioral domains useful for identifying youth with behavioral problems: (a) Intrapersonal Distress, (b) Somatic, (c) Interpersonal Relations, (d) Critical Items, (e) Social Problems, and (f) Behavioral Dysfunction.

The Y-OQ has a 4-week test–retest reliability of .83 and an internal consistency reliability of .97. The concurrent validity of the Y-OQ with the Child Behavior Checklist (CBCL; Achenbach, 1991) and the Conners' Parent Rating Scale (CPRS; Conners, Sitarenios, Parker, & Epstein, 1998) ranges from the .80s to the low .90s. Sensitivity and specificity for distinguishing clinical from healthy community samples are high, and the Y-OQ has been widely accepted for tracking treatment outcome and assessing psychosocial distress (Burlingame et al., 2004).

## Analyses

**Change trajectories.** We used multilevel modeling (MLM; Raudenbush & Bryk, 2002; Singer & Willet, 2003) to evaluate whether Y-OQ scores over time differed between the community and managed care settings (full maximum likelihood, lmer module, lme4 package, R software, Version 2.9.1; Bates & Maechler, 2009; R Development Core Team, 2009). MLM is a form of regression that can be used to predict a participant's score at any particular time (dependent variable) on the basis of a number of independent variables, among which is a *time variable* (e.g., weeks in treatment). MLM is used to estimate the starting point (i.e., intercept) and rate of change during treatment (i.e., slope) for each participant. Some patients received treatment from multiple therapists and at multiple sites and thus were not strictly nested within either. We used R software's lmer modeling to estimate random effects for patient, therapist, and site, controlling for possible correlations within such cross-classified data (see e.g., Raudenbush & Bryk, 2002).

The best-fitting model involved a natural log transformation (base *e*) of weeks in treatment (*LNWEEKS*; based on the deviance and Bayesian information criterion). This transformation models change as curvilinear, occurring quickly at first, then tapering off over time. We used weeks in treatment as the basis for our time variable because it achieved better fit than session number and because precedents in the child treatment literature failed to demonstrate a significant dose–response relationship between sessions attended and treatment outcome (Andrade, Lambert, & Bickman, 2000; Bickman, Andrade, & Lambert, 2002; Salzer, Bickman, & Lambert, 1999).

Our hypothesized model predicted Y-OQ total scores, with a setting variable (1 = *community*, 0 = *managed care*) as a main effect and in interaction with the log of weeks. The main effect for setting tested whether intake Y-OQ scores differed between the community and managed care settings. The interaction between setting and log of weeks tested for differences in rate of change between the two settings. To rule out potential confounds in assessing these differences, we also included the following variables as main effects and in interaction with the log of weeks: yes/no status for having prior treatment, total dose of treatment (i.e., total number of sessions; cf. Baldwin, Berkeljon, Atkins, Olsen, & Nielsen, 2009), total number of Y-OQs (because increased measurement could potentially lead to increased rates of change; see Durham et al., 2002), age, and gender. Our final model included only significant predictors. To facilitate interpretation and reduce multicollinearity, we centered the continuous variables, total number of sessions, total number of Y-OQs, and age around their grand means (e.g., $age - \overline{age}$).

There were other variables for which the community and managed care settings had disparate values. For example, the Y-OQ was administered nearly four times more frequently per month in the managed care setting than in the community setting. If included as covariates in the model, variables such as Y-OQ administrations might have functioned as second (i.e., redundant) indicators of treatment setting, affecting interpretability of the results. We opted to control for these variables instead by estimating additional models for data in which patients were matched one to one between settings, according to the variable values.

In these additional models, separate data sets were used with patients from each setting matched by total weeks in treatment (matched within 1 week), total sessions in treatment (matched within one session), sessions per month (matched within one session), and Y-OQs per month (matched within .17 Y-OQs per month). We also estimated a model using data in which patients were matched one to one by baseline Y-OQ score (within 5 points), thus controlling for the likely correlation between Y-OQ initial score and rate of change. In creating each matched data set, we prioritized exact matches over near matches, drew matches at random from matching candidates (without replacement), and used *t* tests to confirm that means for matching variables did not differ between settings.

**Outcome class.** Whereas the models described focused on how the community and managed care settings differed in terms of Y-OQ change trajectories (i.e., initial scores and rates of change), we also examined how the two settings differed in terms of percentages of cases in various final outcome categories. We determined these outcome classes by comparing overall change scores (i.e., difference between first and last Y-OQ scores) with the Y-OQ's reliable change index of 13 (RCI; Jacobson & Truax, 1991). The RCI is an index of the minimum amount of change in scores that is still distinguishable from measurement error.

The outcome classes we created were as follows: *deterioration* (the final score was at least 13 points worse than baseline), *no reliable change* (the final differed from baseline by less than 13 points), *improvement* (the final score was at least 13 points better than baseline), or *recovery* (the criteria for improvement were met, and the final score was in the subclinical range, i.e., less than 46). Cases whose scores worsened by 13 points or more and remained subclinical at treatment termination were omitted from our comparisons of outcome class percentages between the community and managed care settings (57 cases omitted and 3,954 cases remaining).

We tested a hypothesized model predicting percentages of each outcome class using a setting variable (1 = *community*, 0 = *managed care*). This setting variable was used to test whether the outcome class percentages differed between the community and managed care settings. To rule out potential confounds in assessing these differences, we also included the variables from the hypothesized change trajectory model (with centering for continuous variables), the variables by which we created the matched data subsets, and a centered variable for baseline Y-OQ score. In

iterative revisions to the model, we produced a final outcome class model with nonsignificant parameters removed.

The modeling procedure we used to test whether outcome class percentages differed between the community and managed care settings was partial proportional odds modeling (PPOM; PROC GENMOD, SAS 9.1; Stokes, Davis, & Koch, 2000). In this case, PPOM was used to estimate the odds that patients from each setting would be in one set of outcome classes versus another set (e.g., deterioration and no reliable change vs. improvement and recovery). We structured the model such that it created odds for three comparisons: (a) deterioration versus all other classes, (b) deterioration and no reliable change versus improvement and recovery, and (c) deterioration, no reliable change, and improvement versus recovery. The model produced log odds ratios for each comparison, which, for simplicity, we transformed and reported as odds ratios. We then used these odds to determine the expected percentages of cases in each outcome class for the community and managed care settings. We thus determined whether outcome class percentages differed between settings, while controlling for the effects of other variables.

## Results

### Change Trajectories

Multilevel models showed significant differences in change trajectories for the community and managed care settings. Tables 2A and 2B show estimates for the model that included all the hypothesized covariates. We pared the model down to the best-fitting combination of significant parameters, producing the model estimates shown in Tables 3A and 3B. Figure 1a gives a visual demonstration of how the intercept for the community system trajectory was not significantly different from the intercept for the managed care setting.

The steeper slope (faster rate of change) for the managed care trajectory is also apparent in Figure 1a. For every 1-unit increase in *LNWEEKS* (natural log of weeks in treatment), managed care scores improved 7.7 points, whereas community system scores improved only 3.8 points ($-7.738 + 3.947 = -3.791$; see estimates for slope in Table 3A). These rates represent the improvement in scores after the first 1.7 weeks in treatment (where *LNWEEKS* = 1, weeks = 1.7), given the log transformation equation $LNWEEKS = \log_e(\text{weeks} + 1)$. Note that this amount of improvement requires increasingly longer periods of time as treatment progresses (e.g., where *LNWEEKS* = 2, weeks = 6.4, where *LNWEEKS* = 3, weeks = 19.1), as is expected with the curvilinear trajectory.

Table 2A
*Hypothesized Change Trajectory Model: Fixed Effects*

| | Intercept | | Slope (interaction w/*LNWEEKS*) | |
|---|---|---|---|---|
| Fixed effects | Estimate | *SE* | Estimate | *SE* |
| Intercept[a] | 86.981* | 2.456 | −7.538* | 0.846 |
| Community | −1.464 | 3.204 | 3.445* | 1.083 |
| Prior treatment | −6.844* | 1.281 | 3.059* | 0.477 |
| Total no. of sessions | 0.057 | 0.112 | 0.035 | 0.033 |
| Total no. of Y-OQs | 0.342 | 0.301 | −0.526* | 0.096 |
| Age | −0.539* | 0.163 | −0.021 | 0.061 |
| Girls/young women | −1.745 | 1.143 | −0.186 | 0.423 |

*Note*. *LNWEEKS* = natural log of weeks in treatment; *SE* = standard error; Y-OQs = Youth Outcome Questionnaires.
[a] Estimates for the intercept parameter reflect the mean intercept and slope for the managed care setting—the managed care sample was used as the reference group—where prior treatment = 0 (*no*), girls/young women = 0 (*no*), total sessions = grand mean (centered), total Y-OQs = grand mean (centered), and age = grand mean (centered). Estimates for all other parameters are merely deviations from the intercept constant.
* $p < .05$.

Given that this study's large sample size likely offered statistical power to detect even very small effects, Figure 1a helps to illustrate the size of the effect for setting. For example, after treatment for the grand median of 11 weeks, cases from the different settings would have Y-OQ scores that differed by 9 points (with all other variables held constant). This is a quarter of the size of the standard deviation for Y-OQ baseline scores in this sample (i.e., 36 points), indicating a modest effect.

Further examination of Table 3A shows that estimates for the covariates indicate that intercepts were slightly lower for cases who had received prior treatment or who were older, yet slightly higher for cases with a greater total number of Y-OQs. In addition, rates of change were slower for cases with prior treatment and slightly faster for cases with a greater total number of Y-OQs. Gender had no effect on the trajectories, and age did not affect rate of change.

The predictor variables described helped to explain a portion of the variability in Y-OQ scores. A model with no predictor variables—but one that accounts for the cases' different therapists and treatment sites—facilitates calculation of the variability attributable to individuals, therapists, and sites. From such a model, each random effect estimate is a portion of the overall variability, that is, the sum of the random effects estimates. We calculated such percentages (i.e., intraclass correlation coefficients; Singer & Willet, 2003) and found that 52.6% of the variability occurred between patients, 2.7% occurred between therapists, 7.2% occurred between sites, and 37.5% occurred within patients.

We also examined whether the differences in community versus managed care trajectories would remain when we controlled for additional variables that we figured would be less appropriate as covariates in the model (explained earlier). Table 4 presents estimates for models devised from data in which patients were matched between settings by various characteristics (e.g., baseline Y-OQ score). The corresponding figures, Figures 1b–1f, offer illustrations of setting differences (for cases whose continuous predictor variables had average values and whose dichotomous variables had zero values). The setting effects for intercept remained nonsignificant for all models in the table. All but two of the models in the table continued the trend of significantly slower rates of change (i.e., slopes) for the community setting. The nonsignificant slopes were for the model with patients matched by total number of weeks (almost significant at $p = .057$; shown as though significantly different in Figure 1c) and frequency of Y-OQ administration (Figure 1f), two of the most disparate characteristics between the community and managed care settings in this study (see Table 1A).

Table 2B
*Hypothesized Change Trajectory Model: Random Effects*

| | Intercept | | Slope (*LNWEEKS*) | | |
|---|---|---|---|---|---|
| Random effects | Estimate | *SD* | Estimate | *SD* | Correlation |
| Between clients | 879.463* | 29.656 | 44.454* | 6.667 | −0.281 |
| Between therapists | 3.236* | 1.799 | 0.293* | 0.541 | 0.308 |
| Between sites | 39.377* | 6.275 | 3.855* | 1.963 | 0.829 |
| Within clients (residual) | 409.094* | 20.226 | | | |

*Note*. *LNWEEKS* = natural log of weeks in treatment.
* $p < .05$.

## Outcome Class

The PPOM showed that the community and managed care settings differed in the odds of their cases falling within each outcome class. The odds ratios for constants in Table 5 correspond to managed care (i.e., where community = 0) with default values for the other variables in the model: baseline = grand mean (centered), prior treatment = 0 (*no*), total Y-OQs = grand mean (centered), Y-OQs per month = grand mean (centered), and sessions per month = grand mean (centered). The odds ratios listed in the effects section are multiplicative modifiers of these constants (not additive). For example, the odds ratio for community in the first column of Table 5 indicates that community cases had 1.8 times higher odds of deterioration than managed care cases, with an overall odds of deterioration of 1.800 × .153 = .275. Community's consistent odds ratio modifier of 1.8 for the other outcome class comparisons in the table leads to the more general statement that community cases appeared to have 1.8 times higher odds of less favorable outcomes than managed care cases.

The odds ratios facilitate a calculation of the percentage of cases expected to have an outcome of deterioration. For example, the managed care odds for deterioration would be calculated as .153 ÷ (1 + .153) = .133, indicating that deterioration was the expected outcome 13.3% of managed care cases who had the default values for other variables included in the model. Similar but more complex calculations yielded the expected percentages of each outcome class for both the managed care and community settings, again for cases with default values for other variables included in the model. These modeled percentages were all within 6% of the observed percentages. As the Table 6 shows, expected outcomes were much more positive in the managed care setting, with managed care cases having a 37% higher rate of recovery (modeled) and a 38% lower rate of deterioration (modeled) than community cases. The mean overall change score in the managed care setting showed an improvement of 17.5 points, essentially double the community setting mean of 8.9 points, $t(4009) = -6.80$, $p < .0001$.

Table 5 indicates that in addition to treatment setting, the following covariates influenced the odds of each outcome class: baseline Y-OQ score, prior treatment, total number of Y-OQs, number of Y-OQs per month, and number of sessions per month. It is notable that the differences between settings remained substantial despite the model's control for number of Y-OQs per month and the nonsignificance of total number of weeks; control for each in the change trajectory analyses seemed to diminish the effect of setting.

Table 3A
*Final Change Trajectory Model: Fixed Effects*

| | Intercept | | Slope (interaction w/*LNWEEKS*) | |
|---|---|---|---|---|
| Fixed effects | Estimate | *SE* | Estimate | *SE* |
| Intercept[a] | 86.128* | 2.383 | −7.738* | 0.808 |
| Community | −0.967[b] | 3.006 | 3.947* | 0.992 |
| Prior treatment | −6.844* | 1.281 | 3.097* | 0.476 |
| Total Y-OQs | 0.415[b] | 0.264 | −0.481* | 0.085 |
| Age | −0.603* | 0.144 | | |

*Note*. See Figure 1a. *SE* = standard error; *LNWEEKS* = natural log of weeks in treatment; Y-OQs = Youth Outcome Questionnaires.
[a] Estimates for the intercept parameter reflect the mean intercept and slope for the managed care setting—the managed care sample was used as the reference group—where prior treatment = 0 (*no*), total Y-OQs = grand mean (centered), and age = grand mean (centered). Estimates for all other parameters are merely deviations from the intercept constant. [b] These main effects were retained in the model despite nonsignificance in order for the model to be hierarchically well specified (Peixoto, 1987, 1990).
* $p < .05$.

## Discussion

To our knowledge, this study is the first in which change trajectories and outcome categories have been systematically compared for youths treated in community mental health versus those treated in managed care systems. Two strengths of this study include the use of naturalistic data from two usual care settings in which youths frequently receive services and the use of multilevel modeling to simultaneously examine longitudinal outcomes and account for correlations within therapists and within sites. As hypothesized, individual growth curve models demonstrated that the average rate of change was significantly steeper in the private managed care setting than in the public community mental health setting. Similarly, the PPOM confirmed that negative outcomes (deterioration, no reliable change) were more common in the community mental health setting. In the community setting, 56% of cases were observed to show no significant change or to display a significant increase in symptoms, compared with 46% of cases in the managed care setting. The modeled initial symptom levels were not significantly different across these two usual care settings.

Observed differences in the rate of change and final outcomes across community mental health and managed care settings may be

Table 3B
*Final Change Trajectory Model: Random Effects*

| | Intercept | | Slope (*LNWEEKS*) | | |
|---|---|---|---|---|---|
| Random effects | Estimate | *SD* | Estimate | *SD* | Correlation |
| Between clients | 879.643* | 29.659 | 44.518* | 6.672 | −0.279 |
| Between therapists | 3.857* | 1.964 | 0.290* | 0.539 | 0.309 |
| Between sites | 38.778* | 6.227 | 3.659* | 1.913 | 0.821 |
| Within clients (residual) | 409.093* | 20.226 | | | |

*Note*. *LNWEEKS* = natural log of weeks in treatment.
* $p < .05$.

due to a number of factors. Referral practices and anecdotal reports suggested that youths and families served in the community mental health setting were more often from significantly disadvantaged SES backgrounds; however, SES data from the managed care setting were not available for comparison. Differences in outcomes across settings may also be related to differences in the treatment provided in each setting, although such inferences should be made with caution as only general information was available (obtained through verbal reports from clinicians and supervisors at the two organizations). However, it is interesting to note that although clinicians in both settings reported using eclectic approaches to treatment, the clinicians in the managed care setting reportedly emphasized short-term cognitive–behavioral strategies. Another potentially important difference was in the percentage of clinicians from various professional backgrounds; for example, 73% of therapists in the community mental health setting were identified as social workers, compared with only 27% in the managed care setting.

Results may also be partially explained by differences in the frequency of administration of outcome measures and the methods used in this study to select cases for analysis. Whereas routine practice in the managed care setting was to administer outcome measures each session, measures were administered approximately every four sessions in the community mental health setting. The cases we selected from the original archives had at least two outcome measure administrations; however, this selection process resulted in the exclusion of a significant number of cases from the community mental health sample. Given the spacing between administrations, it is possible that patients in the community setting who improved and discontinued treatment relatively quickly were excluded from our analyses. Consequently, the average rate of change when all community mental health cases are included may be steeper than the trajectory modeled in this study. However, a competing explanation is that more frequent administration of outcome measures in the managed care setting contributed to the steeper rate of change by increasing clinicians' awareness of and response to patient progress or deterioration. In a subset of the data, with patients from each setting matched by frequency of Y-OQ administration, the differences between settings were nonsignificant. However, we are uncertain whether the setting differences became nonsignificant because the model controlled for the influence of Y-OQ administration frequency or simply because the model used a smaller and different sample. This issue warrants further exploration in future research, particularly in light of the relatively modest effect size for setting in the change trajectory analysis with the full sample.

Differing rates of change between settings may also be related to the differing treatment durations in each setting (i.e., three times longer in the community setting). Although total number of sessions did not show significant association with rate of change in this study, Baldwin et al. (2009) found greater numbers of sessions to be linked with slower rates of change in adult patients. In the present study, the difference in rates of change between settings remained when patients from each setting were matched by total number of sessions and again when matched by frequency of sessions. However, the difference in rates of change between settings was not as dramatic when patients were matched by total number of weeks in treatment. Although this may be merely an artifact produced by our use of differing subsamples of data, it may warrant future exploration.

Regardless of the reasons for the observed differences in rate of change and final outcomes across settings, these results have important implications for researchers designing early warning systems to identify patients at risk for treatment failure. Two important components on which many warning systems have been based are the expected change trajectory (the patient's expected symptom level at a given session based on initial symptom level) and the expected rate of treatment failure for a given population. Together these two components are used to calculate the cutoff scores that generate alerts in the warning system. For example, in the system described by Lambert et al. (2002), individual patient progress is compared against empirically derived change trajectories, and an expected treatment failure rate of 10% in adult patients is assumed. Scores at or above the cutoff formed by the top 10% of scores deviating from the expected change trajectory (in the direction of more severe symptoms) triggered an alert. However, in such a system, at-risk cases would likely be underidentified in settings with treatment failure rates significantly higher than 10%. Warning systems used in routine clinical practice are likely to demonstrate greater sensitivity and accuracy when based on treatment failure rates and change trajectories that are typical of the treatment setting in question. The significant differences in treatment failure rates and change trajectories in this study emphasize the need to tailor warning system components and alert criteria to the specific treatment settings in which they will be used.

Considering the clinical implications of this study, the high percentage of cases in each setting with a negative outcome (deterioration or no reliable change) is sobering. In the community mental health setting, 24% of cases showed significantly higher symptom ratings at the last Y-OQ administration compared with when treatment began, and an additional 32% showed no reliable change. In the managed care setting, 14% of cases experienced

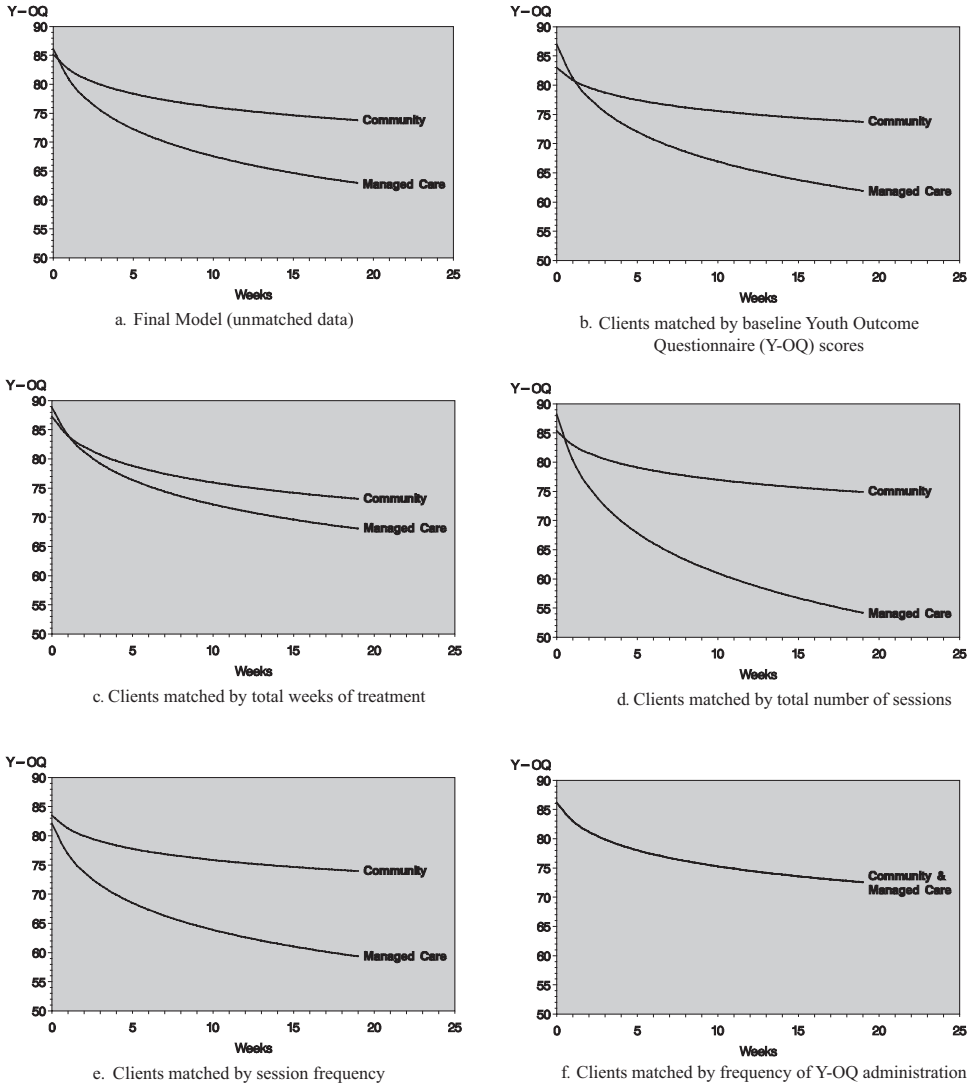


*Figure 1.* Community and managed care trajectories modeled with various data samples. These figures display trajectory differences for cases with average values for continuous predictor variables and zero values for dichotomous predictor variables (variables as in Table 3A). Panel 1a: Final model (unmatched data). Panel 1b: Cases matched by baseline Youth Outcome Questionnaire (Y-OQ) scores. Panel 1c: Cases matched by total weeks of treatment. Panel 1d: Cases matched by total number of sessions. Panel 1e: Cases matched by session frequency. Panel 1f: Cases matched by frequency of Y-OQ administration.

deterioration, and an additional 31% showed no reliable change. Few studies of child and adolescent outcomes in usual care have used reliable change indices to estimate the percentage of cases falling into various outcome categories; however, those existing studies have yielded results comparable to ours. For example, Bybee, Lambert, and Eggett (2007) found a treatment failure rate of 12% in a large sample of children and adolescents served in a managed care system. Also in a managed care setting, Bishop et al. (2005) found a treatment failure rate of 7.2%; however, criteria used to select cases in this study reduced the final sample to 6.6% of the original sample, raising questions about the generalizability of those results. Using a limited subsample of the same archival community mental health data set used in the present study, Warren and colleagues (2009) reported a treatment failure rate of 21%. Overall, results from the present study are consistent with previous findings on the rates of deterioration for children and adolescents in managed care and community mental health settings.

The high percentage of cases with negative outcomes in usual care underscores the need for regular monitoring of youth outcomes in these settings. The American Psychological Association and other professional organizations have emphasized that evidence-based practice includes regular monitoring of patient outcomes such that changes can be made in the approach to treatment if suitable progress is not evident (American Psychological Association Presidential Task Force on Evidence-Based Prac-

Table 4
*Models Based on Matched Data*

| | Match variables | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Baseline Y-OQ ($M$ = 84.7) $n$ = 1,870[a] | | Total weeks ($M$ = 14.1) $n$ = 1,038[b] | | Total sessions ($M$ = 5.9) $n$ = 1,170[c] | | Session frequency ($M$ = 2.2) $n$ = 1,872[d] | | Y-OQ frequency ($M$ = 1.1) $n$ = 1,002[e] | |
| Variable | Estimate | *SE* | Estimate | *SE* | Estimate | *SE* | Estimate | *SE* | Estimate | *SE* |
| Intercept | | | | | | | | | | |
| Intercept | 87.004* | 3.230 | 88.811* | 4.063 | 88.182* | 3.588 | 82.103* | 3.183 | 82.422* | 3.804 |
| Community | −3.972 | 3.807 | −1.583 | 4.830 | −2.822 | 4.304 | 1.371 | 3.753 | 5.165 | 4.451 |
| Slope | | | | | | | | | | |
| Intercept | −8.370* | 0.850 | −6.926* | 0.918 | −11.344* | 1.102 | −7.596* | 0.687 | −4.513* | 0.898 |
| Community | 5.261* | 0.985 | 2.230† | 1.172 | 7.849* | 1.315 | 4.418* | 0.814 | −0.077 | 1.080 |

*Note.* Given that appropriate matches were not available for all clients, samples sizes are listed for each model. This table is intended to be only a summary of how the setting differences remain or vary under various examinations; therefore, estimates for the covariates used in the models (e.g., covariates as in Tables 2 and 3) have been omitted. *SE* = standard error; Y-OQ = Youth Outcome Questionnaire.
[a] See Figure 1b. [b] See Figure 1c. [c] See Figure 1d. [d] See Figure 1e. [e] See Figure 1f.
* $p < .05$. † $p = .057$.

tice, 2006; Institute of Medicine, 2006). Work by Lambert and colleagues (2003) with adult psychotherapy patients provided considerable support for the value of routine monitoring of outcomes, preferably on a session-by-session basis. Among other benefits, regular outcome feedback may be an important defense against the use of potentially harmful therapies (Lilienfeld, 2007). The accuracy of warning systems in child and adolescent populations have been tested in a small number of studies (Bishop et al., 2005; Bybee et al., 2007; Warren et al., 2009); however, an important next step is to conduct randomized clinical trials of feedback and warning systems with these populations. Given that negative outcomes appear to be more common in children and adolescents than in adult populations, the benefits of warning systems and related clinical support tools may be even more pronounced in youth populations than in adults.

Although results of this study revealed important differences in change trajectories and outcomes between the community mental health and managed care settings, the limited nature of the naturalistic archival data did not allow for a systematic examination of the reasons for such differences. Several specific limitations of the data warrant emphasis in interpreting and generalizing this study's results. First, detailed data were unavailable for participants' race and SES, although the latter could be loosely inferred from whether cases were being treated in the community versus managed care setting. For participants treated in the managed care setting, reliable data on diagnosis and number of sessions were lacking. Perhaps results would have differed if any of these variables had been used as covariates or if session number had been used as the time variable for the individual growth models. A second potential limitation, noted previously, is that outcome

Table 5
*Partial Proportional Odds Model of Outcome Class*

| | Deterioration vs. no change, improvement, recovery | | | Deterioration, no change vs. improvement, recovery | | | Deterioration, no change, improvement vs. recovery | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 95% Wald confidence limits | | | 95% Wald confidence limits | | | 95% Wald confidence limits | |
| Variable | Odds ratio | Lower | Upper | Odds ratio | Lower | Upper | Odds ratio | Lower | Upper |
| Constants | 0.153* | 0.138 | 0.170 | 0.789* | 0.728 | 0.855 | 2.940* | 2.640 | 3.274 |
| Effects | | | | | | | | | |
| Community | 1.800* | 1.490 | 2.173 | 1.800* | 1.490 | 2.173 | 1.800* | 1.490 | 2.173 |
| Baseline Y-OQ | 0.992* | 0.990 | 0.993 | 0.992* | 0.990 | 0.993 | 0.992* | 0.990 | 0.993 |
| Prior treatment | 1.349* | 1.185 | 1.536 | 1.349* | 1.185 | 1.536 | 1.349* | 1.185 | 1.536 |
| Total no. of Y-OQs | 0.915* | 0.867 | 0.966 | 0.877 | 0.836 | 0.921 | 0.863* | 0.816 | 0.913 |
| Y-OQs per month | 1.004 | 0.907 | 1.111 | 1.121* | 1.038 | 1.210 | 1.187* | 1.073 | 1.314 |
| Sessions per month | 0.997 | 0.901 | 1.102 | 0.911* | 0.844 | 0.984 | 0.870* | 0.787 | 0.962 |

*Note.* The odds ratios listed in the constants section are those for managed care setting clients (i.e., where community = 0) with default values for the other variables in the model: baseline Youth Outcome Questionnaire (Y-OQ) = grand mean (centered), prior treatment = 0 (*no*), total no. of Y-OQs = grand mean (centered), Y-OQs per month = grand mean (centered), and sessions per month = grand mean (centered). As described in the text, the odds ratios listed in the effects section are multiplicative modifiers of these constants (not additive).
* $p < .05$.

Table 6
*Outcome Classes: Observed and Modeled Percentages*

| Effect/setting | Deterioration (%) | No reliable change (%) | Improvement (%) | Recovery (%) |
|---|---|---|---|---|
| Observed | | | | |
| Community | 24.1 | 31.6 | 27.5 | 16.8 |
| Managed care | 14.3 | 31.4 | 28.5 | 25.8 |
| Modeled | | | | |
| Community | 21.6 | 37.1 | 25.4 | 15.9 |
| Managed care | 13.3 | 30.8 | 30.5 | 25.4 |

measurement was more frequent in the managed care setting (i.e., at nearly every session vs. every fourth session on average in the community setting). More frequent administration of outcome measures in the community mental health setting would allow for more precise estimates of typical change trajectories and final outcomes in this setting. Although we found stable results in a series of models created from simulated data with increasing numbers of measurement occasions omitted—which controlled for analytical response but not for patient or therapist response to added or reduced measurement frequency—there is the potential that our results would have been different had the community data been more complete.

A third limitation to interpreting the results of this study stems from the difference in treatment duration between settings. Given the mean duration of 8 weeks for the managed care setting versus 25 weeks for the community setting, the latter part of predicted trajectories will be less reliable for the managed care setting, where fewer data points inform the predictions. Finally, a fourth limitation is that we used only a single measure of outcome (i.e., no measures other than the Y-OQ); perhaps trajectories and outcomes from the community and managed care settings would compare differently if other outcome measures had been used.

## References


Achenbach, T. M. (1991). *Manual for the Child Behavioral Checklist/4–18 and 1991 profile*. Burlington, VT: University of Vermont, Department of Psychiatry.

Anderson, B. (2007). Collaborative care and motivational interviewing: Improving depression outcomes through patient empowerment interventions. *American Journal of Managed Care, 13,* 103–106.

Andrade, A. R., Lambert, E. W., & Bickman, L. (2000). Dose effect in child psychotherapy: Outcomes associated with negligible treatment. *Journal of the American Academy of Child and Adolescent Psychiatry, 39,* 161–168.

American Psychological Association Presidential Task Force on Evidence-Based Practice. (2006). Evidence-based practice in psychology. *American Psychologist, 61,* 271–285.

Baldwin, S. A., Berkeljon, A., Atkins, D. C., Olsen, J., & Nielsen. S. (2009). Rates of change in naturalistic psychotherapy: Contrasting dose-effect and good-enough level models of change. *Journal of Consulting and Clinical Psychology, 77,* 203–211.

Bates, D., & Maechler, M. (2009). lme4: Linear mixed-effects models using S4 classes (R Package Version 0.999375–31) [Computer software]. Retrieved from http://cran.r-project.org/web/packages/lme4/index.html

Bickman, L., Andrade, A. R., & Lambert, E. W. (2002). Dose response in child and adolescent mental health services. *Mental Health Services Research, 4,* 57–70.

Bishop, M. J., Bybee, T. S., Lambert, M. J., Burlingame, G. M., Wells, M. G., & Poppleton, L. E. (2005). Accuracy of a rationally derived method for identifying treatment failure in children and adolescents. *Journal of Child and Family Studies, 14,* 207–222.

Burlingame, G. M., Cox, J. C., Wells, M. G., Lambert, M. J., Latkowski, M., & Ferre, R. (2005). *The administration and scoring manual of the Youth Outcome Questionnaire.* Salt Lake City, UT: American Professional Credentialing Services.

Burlingame, G. M., Mosier, J. I., Wells, M. G., Atkin, Q. G., Lambert, M. J., Whoolery, M., & Latkowski, M. (2001). Tracking the influence of mental health treatment: The development of the Youth Outcome Questionnaire. *Clinical Psychology and Psychotherapy, 8,* 361–379.

Burlingame, G. M., Wells, A., Lambert, M. J., & Cox, J. (2004). Youth Outcome Questionnaire: Updated psychometric properties. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment*: *Vol. 4* (3rd ed., pp. 235–274). Mahwah, NJ: Erlbaum.

Burns, B. J., Hoagwood, K., & Mrazek, P. J. (1999). Effective treatment for mental disorders in children and adolescents. *Clinical Child and Family Psychology Review, 2,* 199–254.

Bybee, T. S., Lambert, M. J., & Eggett, D. (2007). Curves of expected recovery and their predictive validity for identifying treatment failure. *Dutch Journal of Psychotherapy, 33,* 419–434.

Center for Mental Health Services. (1999). *Annual report to Congress on the evaluation of the Comprehensive Community Mental Health Services for Children and Their Families Program.* Atlanta, GA: ORC Macro.

Conners, C. K., Sitarenios, G., Parker, J. D., & Epstein, J. N. (1998). The revised Conners' Parent Rating Scale (CPRS-R): Factor structure, reliability, and criterion validity. *Journal of Abnormal and Child Psychology, 26,* 257–268.

Durham, C. J., McGrath, L. D., Burlingame, G. M., Schaalje, G. B., Lambert, M. J., & Davies, R. D. (2002). The effects of repeated administrations on self-report and parent-report scales. *Journal of Psychoeducational Assessment, 20,* 240–257.

Finch, A. E., Lambert, M. J., & Schaalje, B. G. (2001). Psychotherapy quality control: The statistical generation of expected recovery curves for integration into an early warning system. *Clinical Psychology and Psychotherapy, 8,* 231–242.

Garland, A. F., Hurlburt, M. S., & Hawley, K. M. (2006). Examining psychotherapy processes in a services research context. *Clinical Psychology: Science and Practice, 13,* 30–46.

Harmon, S. C., Lambert, M. J., Smart, D. M., Hawkins, E., Nielsen, S. L., . . . Lutz, W. (2007). Enhancing outcome for potential treatment failures: Therapist–client feedback and clinical support tools. *Psychotherapy Research, 17,* 379–392.

Hawkins, E. J., Lambert, M. J., Vermeersch, D. A., Slade, K., & Tuttle, K. (2004). The therapeutic effects of providing client progress information to patients and therapists. *Psychotherapy Research, 10,* 308–327.

Howard, K. I., Moras, K., Brill, P. L., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. *American Psychologist, 51,* 1059–1064.

Institute of Medicine. (2006). *Improving the quality of health care for mental and substance-use conditions: Quality Chasm Series.* Washington, DC: National Academies Press.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59,* 12–19.

Kazdin, A. E. (2000). *Psychotherapy for children and adolescents: Directions for research and practice.* New York, NY: Oxford University Press.

Lambert, M. J., & Bergin, A. E. (1994). The effectiveness of psychotherapy. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 143–189). New York, NY: Wiley.

Lambert, M. J., Hansen, N. B., & Finch, A. E. (2001). Patient-focused

research: Using patient outcome data to enhance treatment effects. *Journal of Consulting and Clinical Psychology, 69,* 159–172.

Lambert, M. J., Whipple, J. L., Bishop, M. J., Vermeersch, D. A., Gray, G. V., & Finch, A. E. (2002). Comparison of empirically-derived and rationally-derived methods for identifying patients at risk for treatment failure. *Clinical Psychology and Psychotherapy, 9,* 149–164.

Lambert, M. J., Whipple, J. L., Hawkins, E. J., Vermeersch, D. A., Nielsen, S. L., & Smart, D. W. (2003). Is it time for clinicians to routinely track patient outcome? A meta-analysis. *Clinical Psychology: Science and Practice, 10,* 288–301.

Laurenceau, J. P., Hayes, A. M., & Feldman, G. C. (2007). Some methodological and statistical issues in the study of change processes in psychotherapy. *Clinical Psychology Review, 27,* 682–695.

Lilienfeld, S. O. (2007). Psychological treatments that cause harm. *Perspectives on Psychological Science, 2,* 53–70.

Mohr, D. C. (1995). Negative outcome in psychotherapy: A critical review. *Clinical Psychology: Science and Practice, 2,* 1–27.

National Advisory Mental Health Council. (2001). *Blueprint for change: Research on child and adolescent mental health. A report by the National Advisory Mental Health Council's Workgroup on Child and Adolescent Mental Health Intervention Development and Deployment.* Bethesda. MD: National Institutes of Health/National Institute of Mental Health.

Peixoto, J. L. (1987). Hierarchical variable selection in polynomial regression models. *The American Statistician, 41,* 311–313.

Peixoto, J. L. (1990). A property of well-formulated polynomial regression models. *The American Statistician, 44,* 26–30.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

R Development Core Team. (2009). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Available from http://www.R-project.org

Ringel, J. S., & Sturm, R. (2001). National estimates of mental health utilization and expenditures for children in 1998. *Journal of Behavioral Health Services & Research, 28,* 319–333.

Salzer, M. S., Bickman, L., & Lambert, E. W. (1999). Dose-effect relationship in children's psychotherapy services. *Journal of Consulting and Clinical Psychology, 67,* 228–238.

Silverman, W. K., Kurtines, W. M., & Hoagwood, K. (2004). Research progress on effectiveness, transportability, and dissemination of empirically supported treatments: Integrating theory and research. *Clinical Psychology: Science and Practice, 11,* 295–299.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence.* New York, NY: Oxford University Press.

Stokes, M. E., Davis, C. S., & Koch. G. C. (2000). *Categorical data analysis using the SAS system* (2nd ed.). Cary, NC: SAS Institute.

Warren, J. S., Nelson, P. L., & Burlingame, G. M. (2009). Identifying youth at risk for treatment failure in outpatient community mental health services. *Journal of Child and Family Studies, 18,* 690–701.

Warren, J. S., Nelson, P. L., Burlingame, G. M., & Mondragon, S. A. (2010). *Predicting deterioration in youth psychotherapy: Community mental health vs. managed care settings.* Manuscript submitted for publication.

Weiss, B., Catron, T., Harris, V., & Phung, T. M. (1999). The effectiveness of traditional child psychotherapy. *Journal of Consulting and Clinical Psychology, 67,* 82–94.

Weisz, J. R. (2004). *Psychotherapy for children and adolescents: Evidence-based treatments and case examples.* New York, NY: Cambridge University Press.

Weisz, J. R., Donenberg, G. R., Han, S. S., & Weiss, B. (1995). Bridging the gap between laboratory and clinic in child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology, 63,* 688–701.

Weisz, J. R., & Gray, J. S. (2008). Evidence-based psychotherapy for children and adolescents: Data from the present and a model for the future. *Child and Adolescent Mental Health, 13,* 54–65.

Weisz, J. R., Jensen, A. L., & McLeod, B. D. (2005). Development and dissemination of child and adolescent psychotherapies: Milestones, methods, and a new deployment-focused model. In E. D. Hibbs & P. S. Jensen (Eds.), *Psychosocial treatments for child and adolescent disorders: Empirically based strategies for clinical practice* (2nd ed., pp. 9–39). Washington, DC: American Psychological Association.

Received March 6, 2009
Revision received November 19, 2009
Accepted December 4, 2009 ■