Ohio Mental Health Consumer Outcomes System Report 20: Early Warning System for Identifying Youth at Risk of Treatment Failure

SUMMARY

The purpose of this report is to provide child-serving clinicians with information that can be used to predict which youth will probably end services worse than when they entered services (signal alarm cases), and thus provide an warning signal to alter the course of treatment. Additionally, feedback about outcomes progress in treatment can indicate when youth are making progress in treatment such that termination may be considered. This process is based on comparing the initial and latest Problem Severity scores to critical values related to time in treatment. Research indicates that the number of cases leaving treatment with deteriorated outcomes can be drastically reduced through this process, and that outcomes improvement is increased for others. To gain fuller effect, it is suggested that the Problem Severity scale be completed at each of the first three sessions of care. Though the research indicates that the impact can be gained from providing the clinician with the feedback alone, some suggestions are made about what steps can be taken when an at-risk case is identified. Though this method could be developed to work with Youth Problem Severity scale data, this report refers only to the Parent Problem Severity scale data.

- Using the Parent Problem Severity scale as an early-warning system can identify cases at risk of failure in treatment with 67% sensitivity, specificity of 98%, and an overall hit rate of 95%.
- Measuring Parent Problem Severity during sessions in the first month can add significantly to ability to identify cases at greatest risk of failure.
- Measuring Parent Problem Severity during sessions can identify cases that are making enough progress to suggest that treatment termination may be considered.
- Although at-risk case identification alone is sufficient to improve outcomes, a series of steps are suggested when an at-risk case is identified.
- Outcomes feedback as proposed in this study has also been shown to improve outcomes for not-at-risk cases.
- False positive identification of treatment failure will occur in some cases; however, the cost of false positives, consisting of the extra attention paid to examining the case, is relatively low, and research shows that even these cases benefit from the feedback process and increased attention.
- The benefits highlighted above are maximized by measuring outcomes at each treatment session, but benefit is still accrued even using outcomes on the regular Outcomes System schedule.
- Feedback messages have been built into the change-over-time reports in the Reports Generator.

Prepared by Jim Healy & Kwok Kwan Tam, OPER/ODMH

Introduction

Many sources of data can provide rich feedback to clinicians, and outcomes are one of the most important. Feedback about outcomes can lead to much improved outcomes, which moves outcomes measurement from being "administrative overhead" to an evidence-based clinical practice that improves the clinical outcomes of mental health treatment. Lambert, Hansen & Finch (2001) and Sapyta, Riemer & Bickman (2005) reviewed a technique referred to as patient-focused research, which asks the question "is this particular treatment working for this patient?" When implemented as designed, this method provides clinicians with feedback about the client's status and progress in treatment. In this technique, each treatment modality is tested by measuring progress, and modifications are made to treatment according to the individual's response. In the context of patient-focused research, feedback that indicates a failure to make progress is the primary, but not the only way that feedback has a positive effect in treatment. Feedback that confirms progress in treatment also provides a positive, though weaker impact on outcomes. Sapyta et al. conclude that feedback needs to be fast, specific, objective, and fit the clinician's goals (presumably to improve the outcomes of services) in order to be effective.

Recovery Curves

To make patient-focused research more widely accessible to clinicians, Lambert et al. (2001) drew on over 10,000 cases to establish "recovery curves" that show the amount of progress typically achieved by clients in therapy. These recovery curves were created by using various statistical models to fit the average scale score of the Outcomes Questionnaire-45 (OQ-45) for each session of treatment to a curve. The OQ-45 measures symptom distress, interpersonal relationships, and social role functioning. The recovery curves show the "dose-response" to sessions of therapy for clients who come in at varying levels of severity. In order to provide useful feedback, the researchers used tolerance intervals (NIST-ITL, 2006) to set signal alarm levels. Tolerance intervals estimate the boundaries that contain a given percentage of cases, and are commonly used in process control systems. These tolerance intervals, along with the clinical cutting score, allow for the identification of cases where:

- The client is functioning in the normal range; Recommendation: consider termination. Referred to as White-alarm feedback;
- The rate of change the client is making is in the adequate range; Recommendation: no change in treatment is recommended. Referred to as Green-alarm feedback;
- The rate of change the client is making is less than adequate; Recommendation: consider altering the treatment plan by intensifying treatment, shifting intervention strategies, and monitoring progress especially carefully. Referred to as Yellow-alarm feedback;
- The client is not making the expected level of progress. Chances are he or she may drop out of treatment prematurely, and/or have a negative treatment outcome (getting significantly worse over the course of treatment). Recommendation: steps should be taken to carefully review this case and decide on a new course of action, such as referral for medication, or intensification of treatment. The treatment plan should be reconsidered. Consideration should also be given to presenting this client at a case conference. Referred to as Red-alarm feedback, or signal-alarm case.

Finally, the Lambert et al. study examined the question: does providing the feedback messages above about patient progress on outcomes improve outcomes? To test this, cases were randomly assigned to feedback or no feedback conditions. In all cases, outcomes were administered at every session. In the feedback condition, clinicians were given the feedback message associated with the client's score prior to the start of the session. No other suggestion for treatment modification was given in any way. The no-feedback cases did not receive any information about outcomes. The results showed a powerful impact of feedback: the feedback cases had a 6% rate of reliable deterioration (Jacobson & Truax, 1991, Tam & Healy, 2006), whereas the no-feedback group had a 23% deterioration rate. Additionally, the signal-alarm

cases that received red-feedback were allocated twice the number of sessions as the signal-alarm cases that did not receive feedback. Conversely, the white- and green-level cases that received feedback received less treatment than those white- and green-level cases that received no feedback. The net result was improved outcomes with less treatment.

Finch, Lambert and Schaalje (2001) replicated and extended the feedback system to include 5 levels:

- Blue- Much better than expected progress, beyond the lower 90% tolerance interval;
- White- Better than expected progress, beyond the lower 68% tolerance interval;
- Green- Treatment on course, new score is between the upper and lower 68% tolerance intervals;
- Yellow- Treatment progress is less than expected, beyond the upper 68% tolerance interval;
- Red- Treatment progress is much less than expected, beyond the 90% tolerance level.

The feedback messages would be the same as in the Lambert, Hansen, & Finch study study, with the exception of the Blue-feedback, which could receive a more strongly worded white-feedback message, though the authors do not state a specific blue-feedback message. The authors note that for those cases that begin treatment with a OQ-45 score below the clinical cutting score (i.e., in the normal functioning range), deterioration is expected, and the tolerance intervals around "recovery" curves for individuals with starting scores in this range help identify an expected amount of deterioration versus a truly alarming amount of deviation. The reasons individuals enter treatment with sub-clinical scores can be many, and this is an area that needs further study.

Lambert et al. (2002) compared two methods of identifying cases at risk of treatment failure. The first, described above, is referred to as the "empirically-derived" method based on tolerance intervals. The second was a method based on clinical judgment, referred to as the "rationally-derived" method. The results showed that the empirically-derived method was able to identify 100% of the signal-alarm cases, with 85% being identified in the first three sessions. This compared to an 80% rate for the rationally-derived method, although the rationally derived method identified signal-alarm cases faster, generally after one session. The overall hit rate (number of true negative + number of true positives over all cases) was 84% for the empirically-derived method versus 79% for the rationally-derived method , which was not a statistically significant difference. The empirically-derived method produced false alarms (signal-alarm indication without deterioration) in 20% of cases, and the rationally-derived method did in 18% of cases. Further, an assessment of false alarms showed that these cases did, in fact, have worse outcomes than those cases not identified as signal alarm cases. Seventy-five percent of the false alarm cases failed to make reliable improvement, indicating that even false indication of signal alarm cases may benefit from receiving outcomes feedback.

What matters: Immediate Feedback, Feedback to Consumers or Clinician, or Extensive Clinical Feedback?

Slade et al. (2008) tested the effects of three factors on the impact of feedback:

1) immediate (preceding the session when the outcomes measures were completed) versus one-week delayed feedback,
2) feedback to clinician only versus to the clinician and consumer, and
3) for those consumers who were not on track (red or yellow feedback cases), the effect of getting only a basic feedback message versus getting "clinical support tools" feedback.

Randomly assigned cases to these experimental groups were compared to archival 'treatment as usual' groups who received no feedback. The clinical-support-tools feedback was information about measures of client motivation for therapy, therapeutic alliance, a measure of

social support, and a measure of "perfectionism." This last measure was added as the authors found evidence in the literature that a consumer's sense of perfectionism is a moderating factor for success in treatment. A process for using some of these tools is briefly summarized in the "How to Use this Method" section below. The outcomes feedback groups received weekly feedback reports with a graph showing change over time, a list of critical items, and the appropriate warning messages. The results showed that, although pre- to post-treatment improvement for all conditions was found, feedback significantly improved results over the no-feedback condition. There was no significant difference between cases where the consumer received feedback in addition to clinician. For cases that were not on-track (red or yellow feedback cases), immediate feedback was more effective than delayed feedback. Additionally, the number of sessions varied depending on the experimental group. Overall, all feedback groups received 1.5 sessions more than the no feedback group. For cases not on track, there were 3 more sessions used by groups receiving delayed feedback. Also for not-on-track cases, supplying clinical-support-tools feedback immediately resulted in fewer sessions used. In a separate assessment, the researchers asked consumers in the feedback conditions to rate the benefit of feedback. Ninety-nine percent rated an 8 or higher on a 10-point scale, where 10 was "extremely beneficial", which speaks to the consumer's perception of the relative cost-benefit of completing a 45-item scale every session.

The work of Duncan and Miller at the Institute for the Study of Therapeutic Change parallels that of Lambert and Burlingame. Miller, Duncan, Sorrel and Brown (2005) report on the development of their system of measuring outcomes and therapeutic alliance at each session. Their system finds similar results regarding the identification of potentially failing cases.

Does this same method work in the treatment of children?

Bybee, Lambert and Eggett (2007), developed recovery curves for the Youth Outcomes Questionnaire-30 (YOQ-30) which is used to measure outcomes for youth from 3 to 17 from the youth and parent perspectives. The YOQ-30 taps intrapersonal distress, social problems, behavioral dysfunction, and interpersonal problems. They developed a similar warning system and tested its functioning. They found their system had a sensitivity of 72%, a specificity rate of 90% and an overall hit rate of 88%. These results indicate that the use of this technique for youth is as good as use among adults.

Can outcomes feedback be used outside of psychotherapy?

In a typical psychotherapy model, counseling, sometimes coupled with a pharmacological intervention, are the only services used. Treatment is more or less limited to therapist-client interactions. In our public mental health system, consumers (youth and adult), can receive a variety of services in response to problems that are typically not addressed with a psychotherapy-only approach. All of the work from Lambert and Burlingame, as well as Duncan and Miller, present results in psychotherapy, though their respective systems are in use with those with severe mental illness. However, Ogles et al. (2005) studied the role that feedback and fidelity play in wraparound program models. They provided feedback on outcomes data at two weeks, four weeks, eight weeks and 12 weeks that consisted of the starting and latest outcomes scores. The results indicated few changes attributable to feedback. The most significant result was that those families who received monthly feedback reported a higher rate (69%) of reaching treatment goals than those that did not receive feedback (31%), despite similar gain scores on other measures. Both the feedback and non-feedback groups felt well informed about the progress of treatment. This may have been a result of the much more frequent contact of the family with the treatment team. In personal communication, the lead author indicated that there were few cases that would have met the criteria for being a signal-alarm case, which may have limited certain to-be-expected effects of providing feedback. Also, the feedback itself, a report about starting and last outcomes scores, does not have the same strength of message as the warning signs (red, yellow, green or white messages) used in Lambert & Burlingame's work. The authors theorize that feedback has less impact in services with high frequency of family-treatment team interaction.

Other aspects of the difference between psychotherapy and other treatment modalities regarding the use of feedback to clinicians are also worth noting. The presenting problems that lead people into psychotherapy or other services are different. It may be that feedback will have a differential impact on these various presenting problems. Another difference is the nature of the treatment team. In psychotherapy, there is usually a treatment team of one, and a relatively highly-trained one. In other treatment modalities, it is common that the treatment team is larger (such as a social worker, case manager and psychiatrist), and where the treatment team member in most contact with the family has less training (the case manager).

Other Considerations

Can outcomes that are not collected at every session be used for a warning system? Yes, however the impact will be reduced in relation to the number of points of feedback. Clearly, measures taken at the frequency established in the Outcomes System (initial, three months, six months, annually and at termination) have relatively little opportunity to provide feedback.

Can recovery curves be established based on data that are collected on a set schedule (such as in the Outcomes System) rather than at each treatment session? Yes, in fact, some researchers suggest that it is important to used data from fixed intervals to calculate recovery curves rather than from treatment-dependent intervals, such as sessions, or termination. Feaster, Newman & Rice (2003) note that it may be problematic to use termination data to build recovery curves. They suggest that the data from termination are biased and should not be included in the recovery curve calculation because termination data come from the point in treatment (the end) when it is most likely the case that other non-treatment effects help determine the termination decision. Lambert et al. do not attempt to correct for this, and use termination data in the calculation of the recovery curve.

Methods

The data analysis in this report replicates the methodology from Lambert et al. (2002). From the Ohio Mental Health Consumer Outcomes System, we employed Outcomes assessments from 20,910 individuals who have a valid Ohio Scale Problem Severity assessment within a week from their service admission and at least one valid assessment at 30 days or 90 days[1].

Since the initial score plays a significant role in determining the trajectory of the recovery curve[2] (Finch, Lambert, & Schaalje, 2001), we divided the selected sample into 20 subgroups by percentile of the initial Problem Severity scores. The initial Problem Severity score from each group ranged between two to four points with a larger spread at the two extreme tails. We then constructed the trajectory of the recovery curve for each group.[3]

---

[1] 30 days here refers to Outcomes assessment administered between 16 days and 45 days from service admission. Similarly, 90 days here refers to Outcomes assessment administered between 61 days and 122 days from service admission. If more than one assessment were administered during the above period, the mean of the scores from all assessment were adopted as the score for that period of time.

[2] Since our data were collected on a fixed time schedule rather than the number of session, we are using a modified method from Finch, Lambert, & Schaalje (2001) to establish a linear relationship between the log of the number of weeks and the improvement in Problem Severity.

[3] We adopted the empirically-derived method of Lambert et. al. (2002). After splitting the data into 20 groups, we used PROC MIXED functions of SAS to generate linear models of the recovery curves for each subgroup. For details about the analysis method and procedure, please refer to Finch, Lambert, & Schaalje (2001) & Lambert et al. (2002).

From each of the recovery curves constructed, we derived a 10% tolerance level[4] for each estimated Problem Severity at 30-day assessment. We constructed cut-off scores for the upper and lower bounds for two-tailed one standard deviation for the 30-day estimate from each recovery curve[5]. Using these cut-off scores we identified those individuals that fell above the upper tolerance level at 30-day assessment[6]. These cases were assigned a Red-feedback warning message. In a similar manner, cutoff scores were determined for cases that should receive the Yellow-feedback message. These same intervals are used on the other side of the recovery curve to identify cases that are recovering more rapidly than the norm. These rapidly recovery cases were given the White-feedback and Blue feedback messages. These curves and tolerance levels can be seen in the appendix.

In order to test the effectiveness of the warning system for the therapeutic process, we assessed any significant deterioration from the sample by calculating the difference between 90-day assessment and the initial assessment. Modified from Jacobson and Truax (1991) as constructed in Ohio Mental Health Consumer Outcomes Report#12, any increase of 10 in the Problem Severity[7] between the initial assessment and 90-day assessment was considered as a deteriorating case at 90 days.

---

[4] Tolerance interval indicates the range in which a certain percentage of each individual measurement would fall in the population. This provides a cut-off score to identify cases that fall out of certain range.

[5] The 68% of the sample represented a two-tailed, one standard deviation from the estimated mean.

[6] For those cases with missing 30-day assessments, a mean of a valid initial assessment and a valid 90-day assessment was adopted as the 30-day assessment.

[7] The original Reliable Change from Jacobson and Truax (1991) requires a certain quantitative change in the outcomes as well as the change of clinical status to non-clinical status between the two assessments. In assessing deterioration, we used the reliable change index for outcomes, in our case a 10 point increase in Problem Severity, as a criterion indicating significant deterioration. For details please refer to: http://b9962ed140049a571a710839f1f71c989aaf09ce.gripelements.com/oper/outcomes/reports/rpt.quarterly.12.pdf

Analysis of the predictive validity was undertaken by comparing the signal-alarm cases at 30 days to the deterioration status at 90-days. The results are presented as follows:

	Alarm		No Alarm		Total	
Deteriorate	1428	6.8%	712	3.4%	2140	10.2%
No change/Improve	425	2.0%	18345	87.7%	18770	89.8%
	1853	8.9%	19057	91.1%	20910	
Sensitivity	66.7%					
Specificity	97.7%					
Hit rate	94.6%					
Positive predictive power	77.10%					
Negative predictive power	96.3%					

Sensitivity[8] here refers to the proportion of individuals who show significant deterioration at 90 days among those with a signal alarm. About two-thirds of the sample with alarm signal showed significant deterioration at 90 days.

Specificity[9] refers to the proportion of individuals who do not show deterioration at 90 days among those who do not have a negative signal alarm. In this analysis, 97.7% who showed a negative alarm signal did not have significant deterioration at 90 days.

The hit rate refers to the proportion of correctly identified cases between warning signals and deterioration.

The positive predictive power[10] indicates the percentage of alarm cases that truly deteriorated at 90 days. In this analysis, 77.1% of the cases with a warning signal showed significant deterioration at 90 days.

The negative predictive power[11] shows the percentage of negative alarm cases that did not show significant deterioration at 90 days. In this analysis, 96.3% of the negative alarm cases did not show deterioration.

The above tests indicate that deterioration beyond the critical value at the 30-day assessment is a good predictor of significant deterioration at 90 days.

8 Sensitivity = true positives / (true positives + false negatives)

9 Specificity = true negatives / (true negatives + false positives)

10 Positive predictive value = true positives / (true positives + false positives)

11 Negative predictive value = true negatives / (true negatives + false negatives)

How to Use this Method

Based on our review of the literature and the analysis we have performed, we make the following suggestion for those clinicians who would like to use this method. For youth and families engaged in psychotherapy, measure Problem Severity at the second and third treatment session in addition to the regular Outcomes System administration schedule. This will identify 85% of those cases likely to deteriorate at the end of treatment. We could find no evidence in the literature that feedback would have the same impact for youth and families engaged in other treatment modalities, though we suspect it may. Measurement of Problem Severity at each episode up to termination would likely lead to the other positive gains noted in the literature review. However, this report is based solely on Outcomes data and did not use claims/service data that might have given us evidence of feedback impact on service retention.

The greatest impact of using this method will come from examining the change over time data IMMEDIATELY, before the session where the Outcomes data were collected. A Problem Severity Scale Only version of the Parent instrument is included in the back of this report. If your agency does not have an information system that calculates the score for you, you can do so by summing up the items. If four or fewer items are missing, the easy-to-use "Problem Severity and Functioning Scale Score Look-Up Table" in the appendices will help to calculate the score.

Ideally, access to the warning messages are computerized and provided to staff. If you are using the Data Entry and Reports Template and Reports Generator, you can use the Youth, Parent, Worker Combined Report in the Reports Generator to access this information. You must be using version 4.5 of the Reports Generator or higher (this version has a green background and is listed as version 4.5 in the switchboard form seen when you open the Reports Generator). You can download the Reports Generator at: http://www.mh.state.oh.us/what-we-do/protect-and-monitor/consumer-outcomes/data-flow/template.shtml . If you are not using the Template and Reports Generator, you can use the following manual process.

1. Find the table in the appendix that the Problem Severity starting score falls within.
2. Determine the time in treatment at the point of the latest Problem Severity scores. Find the nearest time in the table. For all ratings falling between assessment and one month interval, use the one month interval.
3. From the nearest time in treatment line, read across the page to find where the range in which the current Problem Severity score falls.
4. Look at the top of the column for the relevant feedback message. The message will be one of the following:

- **Consider Termination:** The client is making much more rapid progress than expected and functioning in the normal range; Recommendation: consider termination.
- **Fast Recovery Noted:** The client is making rapid progress and may be functioning in the normal range; Recommendation: review consumer's need for further treatment.
- **Proceed with Treatment:** The rate of change the client is making is in the adequate range; Recommendation: no change in treatment is recommended.
- **Caution!**: The rate of change the client is making is less than adequate; Recommendation: consider altering the treatment plan by intensifying treatment, shifting intervention strategies, and monitoring progress especially carefully.
- **Warning!**: The client is NOT making the expected level of progress. Chances are he or she may drop out of treatment prematurely, and/or have a negative treatment outcome (getting significantly worse over the course of treatment). Recommendation: Carefully review this case and decide on a new course of action, such as evaluating the need for

medication, or intensification of treatment. Reconsider the treatment plan. Consider presenting this client at a case conference.

Additional Guidance

The improvement in treatment rates discussed in the literature above was obtained primarily from providing feedback messages to clinicians. We believe that with early warning, most clinicians know what they need to do to make a course correction in treatment. For those who would like use the "clinical support tool" approach to follow up on not-on-track cases (Warning! and Caution! feedback cases), there is a large amount of literature available about what variables make treatment effective and what interventions could be applied to failing cases. A full review of this is beyond the scope of this report. Thankfully, however, Harmon et al. (2005) offer a simple stepwise process and handbook on what to consider whenever a clinician receives a Yellow or Red feedback message, and is well worth the trouble to track down a copy if you intend to use this process. Here is summary of the stepwise process, along with some alternative resources if you cannot get the article:

1. Assess the therapeutic alliance, preferably with a standard measure:
   - Duncan and Miller's Session Rating Scale (SRS) http://www.talkingcure.com/index.asp?id=106
   - Penn Helping Alliance questionnaire (HAq) http://www.med.upenn.edu/cpr/instruments.html
   - Horvath and Greenberg's Working Alliance Inventory (WAI) http://www.mps.mb.ca/Continuing%20Ed/Scales/WAIclient.html
   - California Psychotherapy Alliance Scales (CALPAS)

2. Assess the youth and parents/family readiness to change. We could not find any measures of readiness to change specifically for youth and/or families in mental health treatment. However, the transtheoretical model (or stages of change model), associated with Prochaska, Norcross, and DiClementi, and motivational interviewing, associated with Miller and Rollnick, both provide conceptual models for what to do. We liked these sources, but many are available on the net:
   - Stages of Changes: http://en.wikipedia.org/wiki/Transtheoretical_Model
   - Motivational Interviewing: http://www.motivationalinterview.org/

3. Assess the person's social support. The person may lack positive support, or may have support systems that reinforce unhelpful behavior.
   - Multidimensional Scale of Perceived Social Support. A 12-item scale that taps support from family, friends and significant others. Available in the back of this article: http://www.psychosomaticmedicine.org/cgi/reprint/49/4/331 .
     Items 3, 4, 8 and 11 measures family support; items 1, 2, 5, 6, 7, 9 & 12 measure friend support; and item 10 measures significant other support.

4. Reassess the diagnostic formulation, and reconsider if an effective treatment option has not been attempted.

5. Assess whether medication is an appropriate action.

6. Assess the consumer's tendency toward perfectionism:
   - There are many sites that discuss perfectionism; we liked this one: http://www.apa.org/monitor/nov03/manyfaces.html
   - Perfectionism Inventory (PI), an 8-scale inventory: http://www.psych.appstate.edu/faculty/hill/Hill%20-%20Perfectionism%20Inventory.pdf

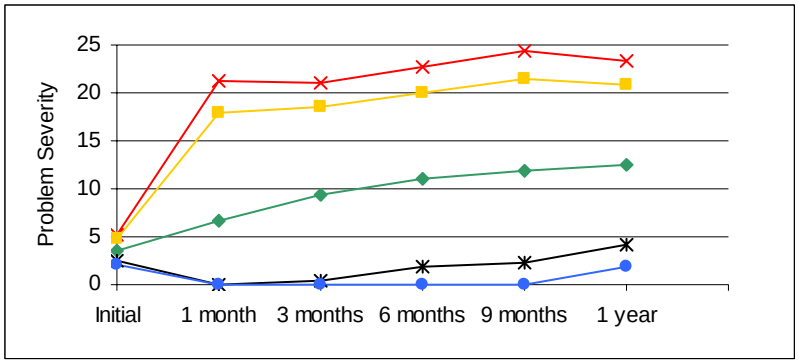
Harmon et al. offer few suggestions about steps 4, 5, or 6, and we are unwilling to offer alternate resources for these steps.

Expected Progress Tables and Graphs

Initial Parent Problem Severity = 1-5

	Red >=90%	Yellow <90% & >=84%	Green <84% & >16%	White <=16% & >10%	Blue <=10%
1 month	21	18	7	0*	0*
3 months	21	18	9	0	0*
6 months	23	20	11	2	0*
9 months	24	22	12	2	0*
1 year	23	21	13	4	2

* cut-off scores were set to zero for negative estimations.

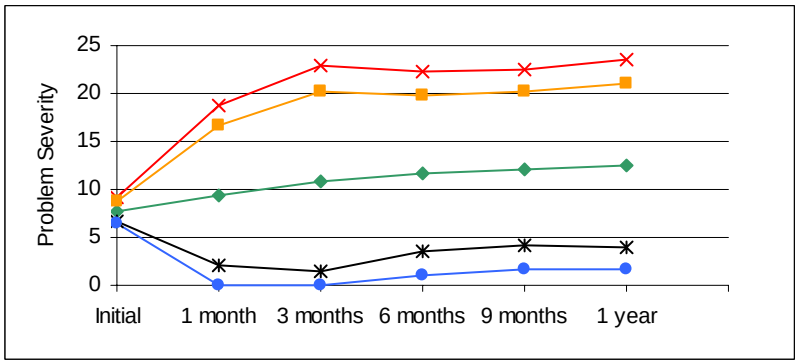


Note: Higher Problem Severity score indicates higher level of problem.

Initial Parent Problem Severity = 6-9

	Red >=90%	Yellow <90% & >=84%	Green <84% & >16%	White <=16% & >10%	Blue <=10%
1 month	19	17	9	2	0
3 months	23	20	11	1	0*
6 months	22	20	12	3	1
9 months	23	20	12	4	2
1 year	23	21	13	4	2

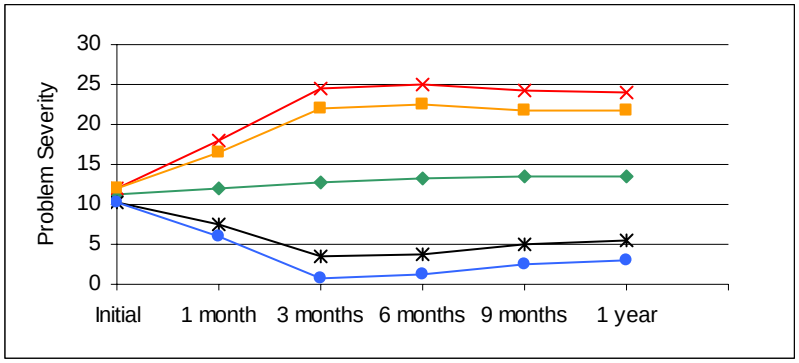
* cut-off score was set to zero for negative estimation.



Note: Higher Parent Problem Severity score indicates higher level of problem.

Initial Parent Problem Severity = 10-12

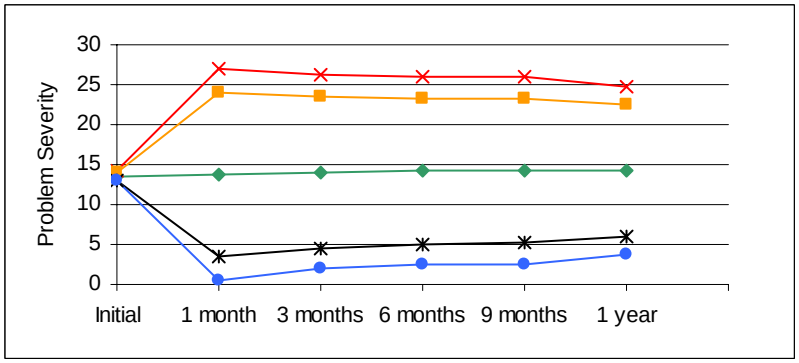
	Red >=90%	Yellow <90% & >=84%	Green <84% & >16%	White <=16% & >10%	Blue <=10%
1 month	18	17	12	7	6
3 months	25	22	13	4	1
6 months	25	22	13	4	1
9 months	24	22	13	5	2
1 year	24	22	14	5	3



Note: Higher Parent Problem Severity score indicates higher level of problem.

Initial Parent Problem Severity = 12-14

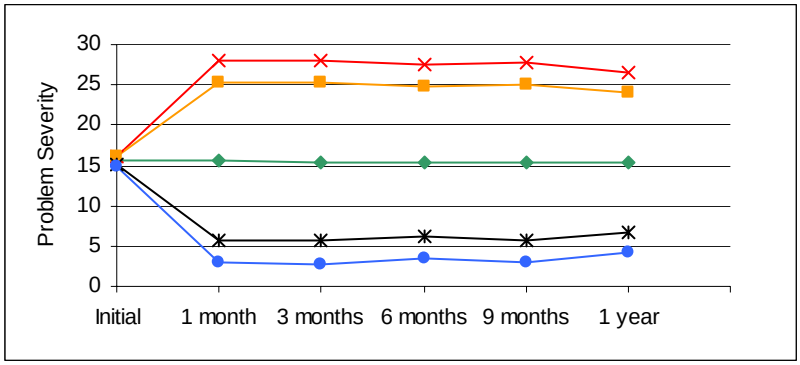
	Red >=90%	Yellow <90% & >=84%	Green <84% & >16%	White <=16% & >10%	Blue <=10%
1 month	27	24	14	4	1
3 months	26	23	14	5	2
6 months	26	23	14	5	2
9 months	26	23	14	5	3
1 year	25	22	14	6	4



Note: Higher Parent Problem Severity score indicates higher level of problem.

Initial Parent Problem Severity = 15-16

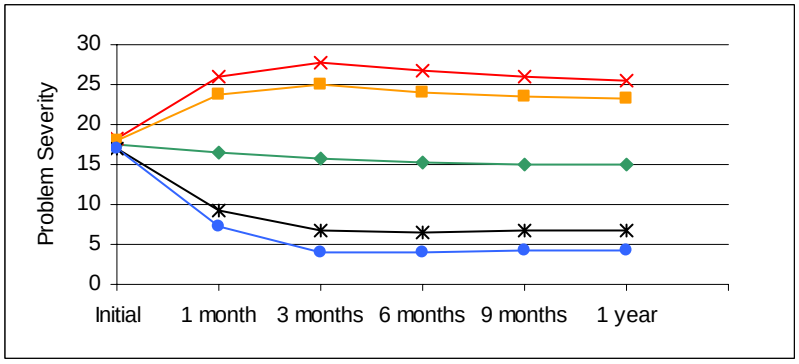
	Red >=90%	Yellow <90% & >=84%	Green <84% & >16%	White <=16% & >10%	Blue <=10%
1 month	28	25	16	6	3
3 months	28	25	15	6	3
6 months	27	25	15	6	3
9 months	28	25	15	6	3
1 year	27	24	15	7	4



Note: Higher Parent Problem Severity score indicates higher level of problem.

Initial Parent Problem Severity = 17-18

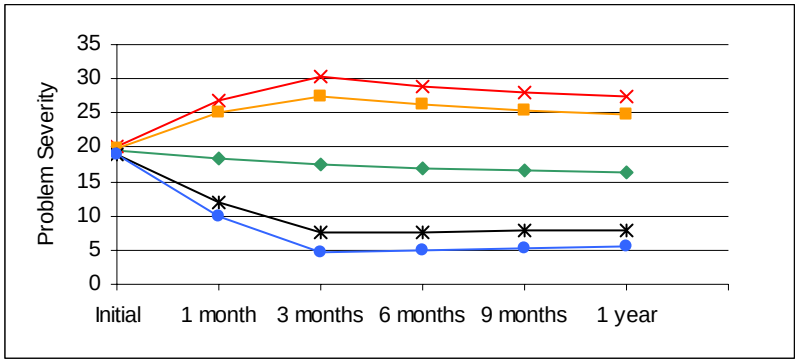
	Red >=90%	Yellow <90% & >=84%	Green <84% & >16%	White <=16% & >10%	Blue <=10%
1 month	26	24	17	9	7
3 months	28	25	16	7	4
6 months	27	24	15	7	4
9 months	26	24	15	7	4
1 year	26	23	15	7	4



Note: Higher Parent Problem Severity score indicates higher level of problem.

Initial Parent Problem Severity = 19-20

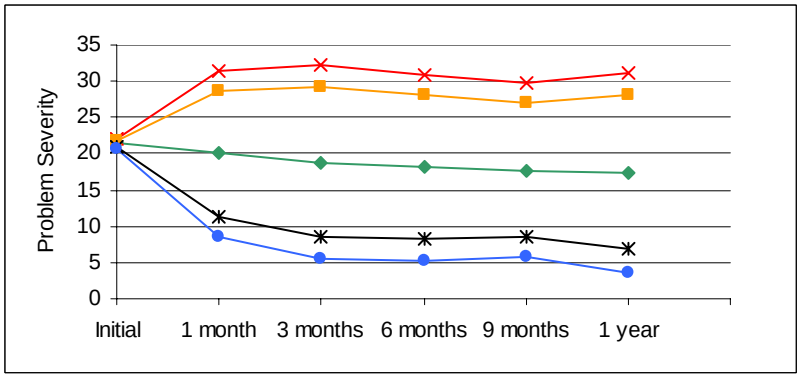
	Red >=90%	Yellow <90% & >=84%	Green <84% & >16%	White <=16% & >10%	Blue <=10%
1 month	27	25	18	12	10
3 months	30	27	17	8	5
6 months	29	26	17	8	5
9 months	28	25	17	8	5
1 year	27	25	16	8	6



Note: Higher Parent Problem Severity score indicates higher level of problem.

Initial Parent Problem Severity = 21-22

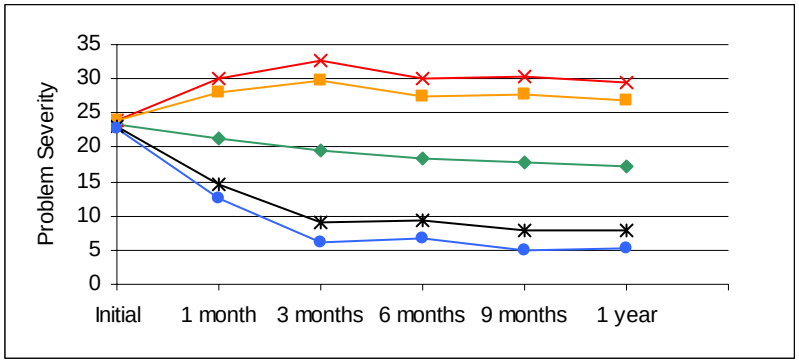
	Red >=90%	Yellow <90% & >=84%	Green <84% & >16%	White <=16% & >10%	Blue <=10%
1 month	31	29	20	11	9
3 months	32	29	19	8	6
6 months	31	28	18	8	5
9 months	30	27	18	8	6
1 year	31	28	17	7	4



Note: Higher Parent Problem Severity score indicates higher level of problem.

Initial Parent Problem Severity = 23-24

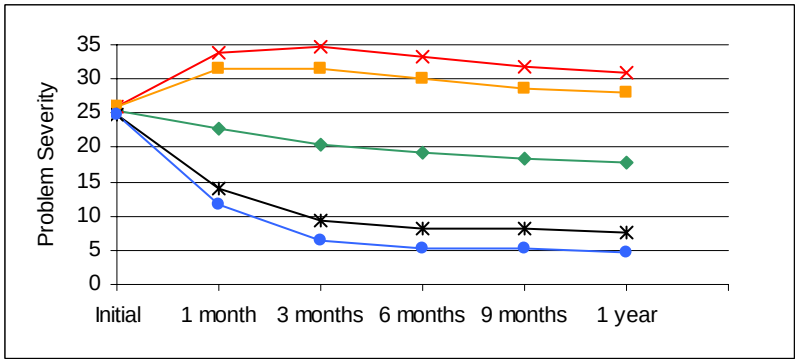
	Red >=90%	Yellow <90% & >=84%	Green <84% & >16%	White <=16% & >10%	Blue <=10%
1 month	30	28	21	15	13
3 months	33	30	19	9	6
6 months	30	28	18	9	7
9 months	30	28	18	8	5
1 year	29	27	17	8	5



Note: Higher Parent Problem Severity score indicates higher level of problem.

Initial Parent Problem Severity = 25-26

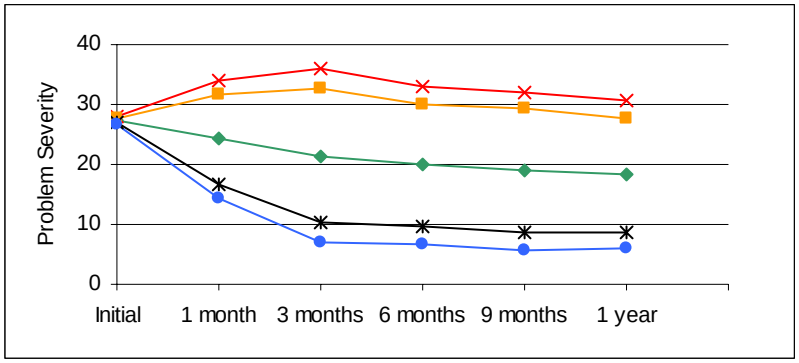
	Red >=90%	Yellow <90% & >=84%	Green <84% & >16%	White <=16% & >10%	Blue <=10%
1 month	34	31	23	14	12
3 months	35	32	21	9	6
6 months	33	30	19	8	5
9 months	32	29	18	8	5
1 year	31	28	18	8	5



Note: Higher Parent Problem Severity score indicates higher level of problem.

Initial Parent Problem Severity = 27-28

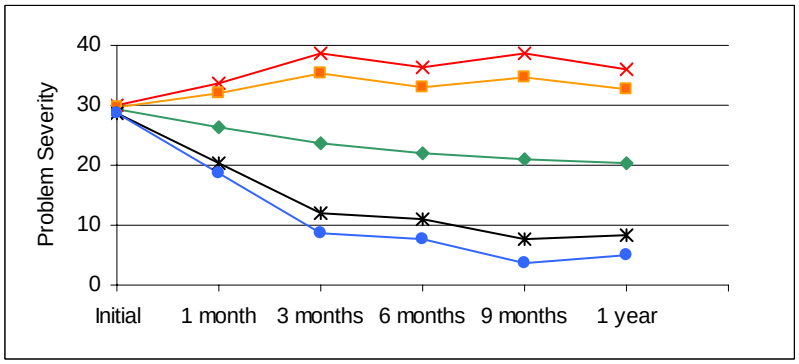
	Red >=90%	Yellow <90% & >=84%	Green <84% & >16%	White <=16% & >10%	Blue <=10%
1 month	34	32	24	17	14
3 months	36	33	21	10	7
6 months	33	30	20	10	7
9 months	32	29	19	9	6
1 year	31	28	18	9	6



Note: Higher Parent Problem Severity score indicates higher level of problem.

Initial Parent Problem Severity = 29-30

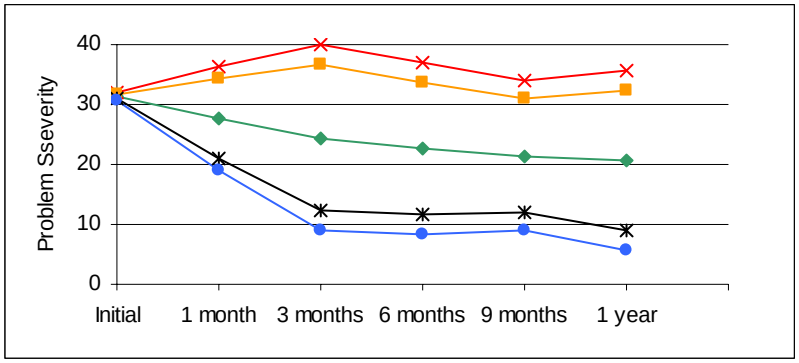
	Red >=90%	Yellow <90% & >=84%	Green <84% & >16%	White <=16% & >10%	Blue <=10%
1 month	34	32	26	20	19
3 months	39	35	24	12	9
6 months	36	33	22	11	8
9 months	39	35	21	8	4
1 year	36	33	20	8	5



Note: Higher Parent Problem Severity score indicates higher level of problem.

Initial Parent Problem Severity = 31-32

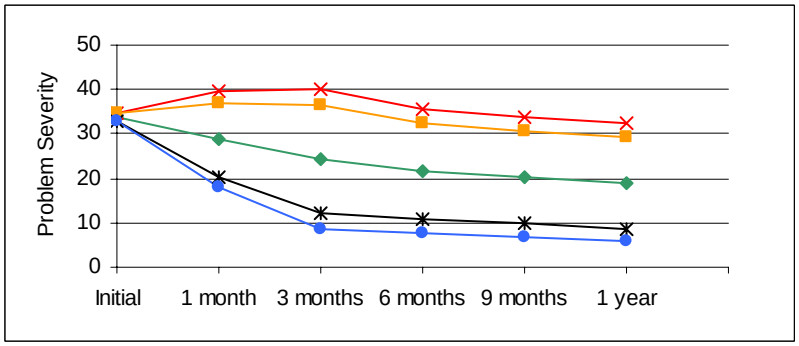
	Red >=90%	Yellow <90% & >=84%	Green <84% & >16%	White <=16% & >10%	Blue <=10%
1 month	36	34	28	21	19
3 months	40	37	24	12	9
6 months	37	34	23	12	8
9 months	34	31	21	12	9
1 year	36	32	21	9	6



Note: Higher Parent Problem Severity score indicates higher level of problem.

Initial Parent Problem Severity = 33-35

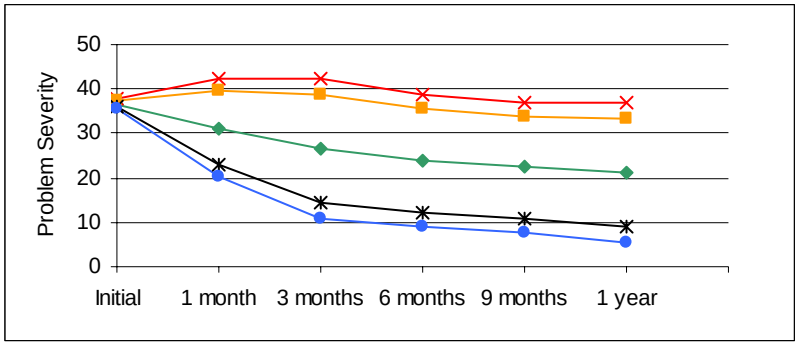
	Red >=90%	Yellow <90% & >=84%	Green <84% & >16%	White <=16% & >10%	Blue <=10%
1 month	39	37	29	20	18
3 months	40	37	24	12	8
6 months	36	33	22	11	8
9 months	34	31	20	10	7
1 year	32	29	19	9	6



Note: Higher Parent Problem Severity score indicates higher level of problem.

Initial Parent Problem Severity = 36-38

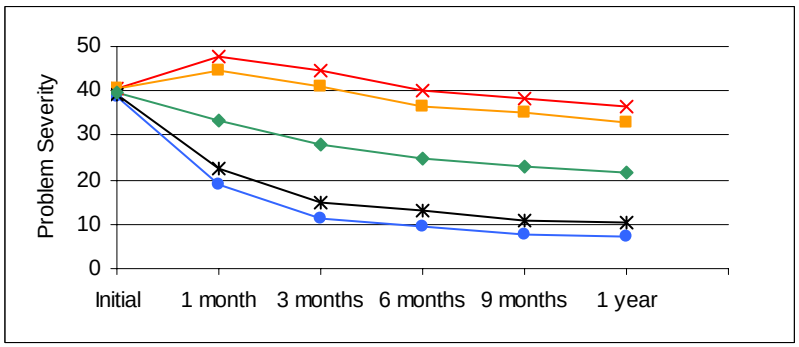
	Red >=90%	Yellow <90% & >=84%	Green <84% & >16%	White <=16% & >10%	Blue <=10%
1 month	42	40	31	23	20
3 months	42	39	27	14	11
6 months	39	35	24	12	9
9 months	37	34	22	11	7
1 year	37	33	21	9	5



Note: Higher Parent Problem Severity score indicates higher level of problem.

Initial Parent Problem Severity = 39-41

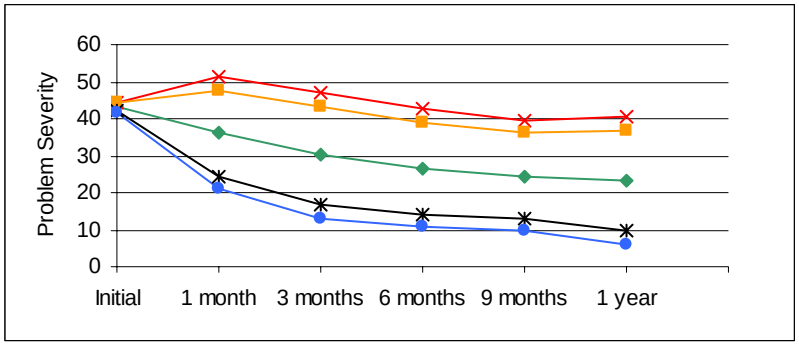
	Red >=90%	Yellow <90% & >=84%	Green <84% & >16%	White <=16% & >10%	Blue <=10%
1 month	48	45	33	22	19
3 months	45	41	28	15	11
6 months	40	37	25	13	10
9 months	38	35	23	11	8
1 year	36	33	22	10	7



Note: Higher Parent Problem Severity score indicates higher level of problem.

Initial Parent Problem Severity = 42-45

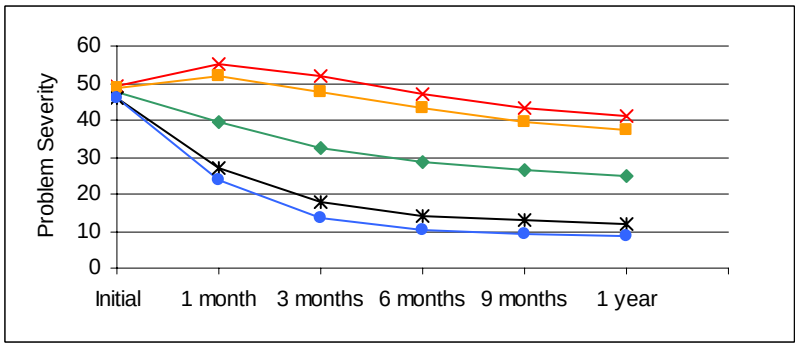
	Red >=90%	Yellow <90% & >=84%	Green <84% & >16%	White <=16% & >10%	Blue <=10%
1 month	51	48	36	24	21
3 months	47	43	30	17	13
6 months	42	39	27	14	11
9 months	39	36	25	13	10
1 year	40	37	23	10	6



Note: Higher Parent Problem Severity score indicates higher level of problem.

Initial Parent Problem Severity = 46-50

	Red >=90%	Yellow <90% & >=84%	Green <84% & >16%	White <=16% & >10%	Blue <=10%
1 month	55	52	39	27	24
3 months	52	47	33	18	13
6 months	47	43	29	14	10
9 months	43	39	26	13	9
1 year	41	37	25	12	8

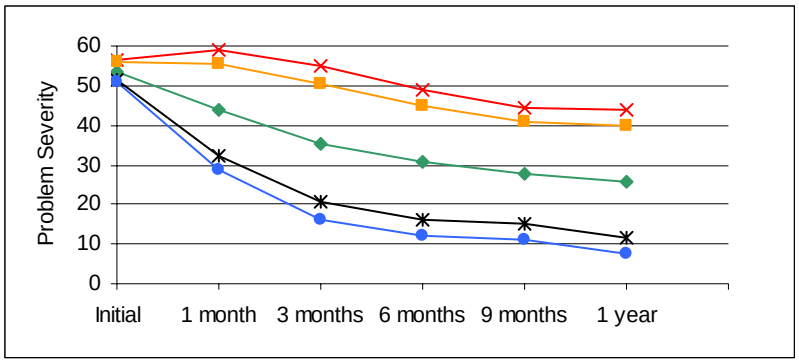


Note: Higher Parent Problem Severity score indicates higher level of problem.

Initial Parent Problem Severity = 51-58

	Red	Yellow	Green	White	Blue
	>=90%	<90% & >=84%	<84% & >16%	<=16% & >10%	<=10%
1 month	59	56	44	32*	29*
3 months	55	51	36	20	16
6 months	49	45	31	16	12
9 months	44	41	28	15	11
1 year	44	40	26	12	8

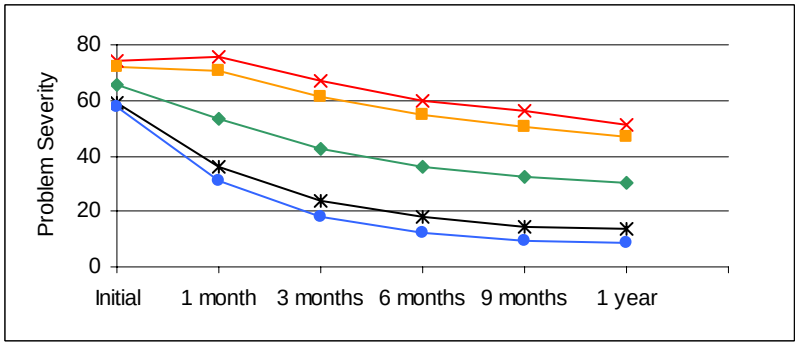
* Scores not in normal range



Note: Higher Parent Problem Severity score indicates higher level of problem.

Initial Parent Problem Severity >=59

	Red	Yellow	Green	White	Blue
	>=90%	<90% & >=84%	<84% & >16%	<=16% & >10%	<=10%
1 month	76	71	53	36*	31*
3 months	67	61	43	24	18
6 months	60	55	36	18	13
9 months	56	51	33	14	9
1 year	51	47	30	13	9



Note: Higher Parent Problem Severity score indicates higher level of problem.

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

Parent Rating – Problem Severity Scale Only

P

Child's Name: _____ Date: _____ Child's Grade: ___ ID#: _____ Completed by Agency

Child's Date of Birth: _____ Child's Sex: ❑ Male ❑ Female Child's Race: _____

Form Completed By: ❑ Mother ❑ Father ❑ Step-mother ❑ Step-father ❑ Other: _____

Instructions: Please rate the degree to which your child has experienced the following problems in the past 30 days.	Not at All	Once or Twice	Several Times	Often	Most of the Time	All of the Time
1. Arguing with others	0	1	2	3	4	5
2. Getting into fights	0	1	2	3	4	5
3. Yelling, swearing, or screaming at others	0	1	2	3	4	5
4. Fits of anger	0	1	2	3	4	5
5. Refusing to do things teachers or parents ask	0	1	2	3	4	5
6. Causing trouble for no reason	0	1	2	3	4	5
7. Using drugs or alcohol	0	1	2	3	4	5
8. Breaking rules or breaking the law (out past curfew, stealing)	0	1	2	3	4	5
9. Skipping school or classes	0	1	2	3	4	5
10. Lying	0	1	2	3	4	5
11. Can't seem to sit still, having too much energy	0	1	2	3	4	5
12. Hurting self (cutting or scratching self, taking pills)	0	1	2	3	4	5
13. Talking or thinking about death	0	1	2	3	4	5
14. Feeling worthless or useless	0	1	2	3	4	5
15. Feeling lonely and having no friends	0	1	2	3	4	5
16. Feeling anxious or fearful	0	1	2	3	4	5
17. Worrying that something bad is going to happen	0	1	2	3	4	5
18. Feeling sad or depressed	0	1	2	3	4	5
19. Nightmares	0	1	2	3	4	5
20. Eating problems	0	1	2	3	4	5

(Add ratings together) Total _____

Problem Severity and Functioning Scale Score Look-Up Table

Directions: Use the tables below to look up the value of the Problem Severity or Functioning scale score when you have missing values. Add up the values of the completed items. If the number of missing items is less than five, then find the sum of the non-missing items in one of the columns under the arrow, then go across to find the number of missing items that are listed across the top of the table. The value at the intersection of the sum of completed items and the number of missing items is the scale score. If more than four items are missing, no scale score is calculated.

Sum of completed items ↓	Number of Missing Items			
	1	2	3	4
0	0	0	0	0
1	1	1	1	1
2	2	2	2	3
3	3	3	4	4
4	4	4	5	5
5	5	6	6	6
6	6	7	7	8
7	7	8	8	9
8	8	9	9	10
9	9	10	11	11
10	11	11	12	13
11	12	12	13	14
12	13	13	14	15
13	14	14	15	16
14	15	16	16	18
15	16	17	18	19
16	17	18	19	20
17	18	19	20	21
18	19	20	21	23
19	20	21	22	24
20	21	22	24	25
21	22	23	25	26
22	23	24	26	28
23	24	26	27	29
24	25	27	28	30
25	26	28	29	31
26	27	29	31	33
27	28	30	32	34
28	29	31	33	35
29	31	32	34	36
30	32	33	35	38
31	33	34	36	39
32	34	36	38	40
33	35	37	39	41
34	36	38	40	43
35	37	39	41	44
36	38	40	42	45
37	39	41	44	46
38	40	42	45	48
39	41	43	46	49
40	42	44	47	50
41	43	46	48	51
42	44	47	49	53
43	45	48	51	54
44	46	49	52	55
45	47	50	53	56
46	48	51	54	58
47	49	52	55	59
48	51	53	56	60
49	52	54	58	61
50	53	56	59	63

Sum of completed items ↓	Number of Missing Items			
	1	2	3	4
51	54	57	60	64
52	55	58	61	65
53	56	59	62	66
54	57	60	64	68
55	58	61	65	69
56	59	62	66	70
57	60	63	67	71
58	61	64	68	73
59	62	66	69	74
60	63	67	71	75
61	64	68	72	76
62	65	69	73	78
63	66	70	74	79
64	67	71	75	80
65	68	72	76	81
66	69	73	78	83
67	71	74	79	84
68	72	76	80	85
69	73	77	81	86
70	74	78	82	88
71	75	79	84	89
72	76	80	85	90
73	77	81	86	91
74	78	82	87	93
75	79	83	88	94
76	80	84	89	95
77	81	86	91	96
78	82	87	92	98
79	83	88	93	99
80	84	89	94	100
81	85	90	95	
82	86	91	96	
83	87	92	98	
84	88	93	99	
85	89	94	100	
86	91	96		
87	92	97		
88	93	98		
89	94	99		
90	95	100		
91	96			
92	97			
93	98			
94	99			
95	100			

References

Claiborn, C.D., & Goodyear, R.K. (2005). Feedback in psychotherapy. Journal of Clinical Psychology, 61 (2), 209-217.

Bybee, T.S., Lambert, M.J., Eggett, D. (2007) Curves of expected recovery and their predictive validity for identifying treatment failure. Psychotherapie*, 33(6)*,

Feaster, D. J., Newman, F.L. & Rice, C. (2003). Longitudinal analysis when the experimenter does not determine when treatment ends: what is dose response? Clinical Psychology and Psychotherapy, 10, 352-360.

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.

Harmon, C., Hawkins, E.J., Lambert, M.J, Slade, K. & Whipple, J.L. (2005) Improving outcomes for poorly responding clients: the use of clinical support tools and feedback to clients. Journal of Clinical Psychology, 61(2), 175-185.

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(1): 12-19.

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 (2), 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. Journal of Consulting and Clinical Psychology, 9, 149-164.

Miller, S.D., Duncan, B.L., Sorrell, R., & Brown, G.S. (2005).The Partners in Change Outcome Management System. Journal of Clinical Psychology, 61(2),199-208.

NIST-ITL. What intervals contain a fixed percentage of the population values? (2006). In National Institute of Standards and Technology, Information Technology Laboratory Engineering statistics handbook, Retrieved October, 2008, from http://www.itl.nist.gov/div898/handbook/prc/section2/prc264.htm

Ogles, B.M., Carlston, D., Hatfield, D., Melendez, G., Dowell, K., & Fields, S.A. (2006) The role of fidelity and feedback in the wraparound approach. Journal of Child and Family Studies. 15(1), 115-129.

Sapyta, J., Riemer, M., & Bickman, L. (2005). Feedback to clinicians: theory, research, and practice. Journal of Clinical Psychology, 61 (2), 145-153.

Slade, K., Lambert, M.J., Harmon, S.C., Smart, D.W. & Bailey, R. (2008). Improving psychotherapy outcome: the use of immediate electronic feedback and revised clinical support tools. Clinical Psychology and Psychotherapy,15, 287-303.