



Ohio Mental Health Consumer Outcomes System

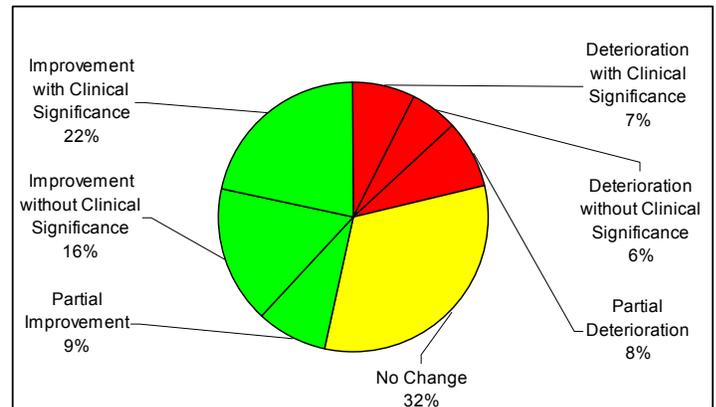
Report 16: *Benchmarks for Reliable Change and Clinical Significance on the Ohio Scales for Youth - Problem Severity and Functioning Scales*

The purpose of this report is to provide benchmarks for improvement and Clinical Significance in comparing treatment outcomes between two time points for the Problem Severity and Functioning scales of the Ohio Scales for Youth. Using the concepts of Reliable Change and Clinical Significance, we developed seven categories in assessing changes over time. In this report, we present the distributions of improvement in the two scales over four pair-wise comparisons: 1) Initial and 3-month assessments, 2) Initial and 6-month assessments, 3) Initial and 1-year assessments, and 4) Initial and Termination assessments

Benchmarks are presented overall, and by age, race, and gender. Additionally, benchmarks are presented by intake score, county size, and primary diagnostic group where these factors are significant.

SUMMARY

- For Problem Severity, about 21% of youth with completed Outcomes measures achieve a Reliable Change with Clinical Significance (move from the clinical to the non-clinical range) at 3 months, 24% at 6 months, 25% at 1 year, and 28% at termination. When the other improvement groups are added, a total of about 47% show some improvement at 3 months, 49% at 6 months, 50% at 1 year, and 52% at termination. As an example, the pie chart to the right shows the results for the Youth Measure from Initial to 3 months for 10,111 youth. The improvement categories are in green on the left side, no change in yellow at the lower right, and deterioration categories in red at the upper right.



- For Functioning, about 17% of youth with completed measures achieve a Reliable Change with Clinical Significance (move from the clinical to the non-clinical range) at 3 months, 20% at 6 months, 22% at 1 year, and 23% at termination. When the other improvement groups are added, a total of about 46% show some improvement at 3 months, 49% at 6 months, 51% at 1 year, and 51% at termination.
- Results from analysis of variance show that initial score in the clinical range, age at intake, and county size are significant factors predicting the improvement in Problem Severity and Functioning scores.
- In many comparisons, youth ages 13 to 15 show a higher percentage in improvement in both Problem Severity and Functioning, while youth aged 5 to 8 show lower improvement rates.
- In most comparisons, youth with Mood Disorders experience a higher rate of improvement in Problem Severity.
- Youth in small and medium counties show slightly more improvement in Problem Severity and Functioning than those in large counties.
- Youth with intake Outcomes scores (Problem Severity or Functioning) in the clinical range¹ consistently show greater improvement in Outcomes scores compared to youth who begin treatment in the non-clinical range.

¹ The clinical cutoff for the Problem Severity score of the Ohio Scales is 20 for all three ratings. Any scores 20 or above are classified as within the clinical range, while any scores below 20 will be treated as falling into the non-clinical range. For the Functioning score, the clinical cutoff is 50 for parent and worker ratings, and 60 for youth ratings. For details please refer to Ogles & Healy (2005).

Table of Contents

SUMMARY	1
What are Reliable Change and Clinical Significance?	3
Seven categories of changes in this report	3
Multivariate Analysis: Exploring factors predicting changes	7
Discussion	7
Understanding the charts	8
Change in Ohio Scales – Problem Severity	9
Problem Severity—Overall Changes	9
Changes by Gender—Problem Severity	13
Changes by Age—Problem Severity	17
Changes by Race—Problem Severity	21
Changes by Diagnostic Group—Problem Severity	25
Changes by County Size—Problem Severity	29
Changes by Intake Problem Severity—Problem Severity	33
Change in Ohio Scales - Functioning	37
Overall Changes—Functioning	37
Changes by Gender—Functioning	41
Changes by Age—Functioning	45
Changes by Race—Functioning	49
Changes by Diagnostic Group—Functioning	53
Changes by County Size—Functioning	57
Changes by Intake Functioning Level—Functioning	61
References	65
Appendix A—ANOVA Results Tables	66

What are Reliable Change and Clinical Significance?

Various research efforts have struggled with defining what constitutes improvement in psychotherapy, and most of these researchers focus on the concept of *Reliable and Clinically Significant Change* (some examples: Jacobson & Truax, 1991; Lunnen & Ogles, 1998; McGlinchey, Atkins, & Jacobson, 2002). The broad concept of Reliable and Clinically Significant Change² can be addressed in two parts:

Reliable Change - Is the change of sufficient magnitude to be confident that the change is beyond that which could be attributed to measurement error? An assessment of Reliable Change is a valuable benchmark indicating that progress in treatment has occurred. *Reliable Change* (RC) is based upon the reliability or consistency of the measurement instrument. A reliable change is usually assessed as the difference between pretest and posttest scores against a certain critical level. When the difference between two time points is above the critical level, the difference is not likely to be an artifact of measurement error.

Clinical Significance – How does the end state of the client compare with the scores observed in socially and clinically meaningful comparison groups? In this report, we adopted *Clinical Significance* as a return to non-clinical functioning from a clinical population (Jacobson & Truax, 1991), using the reliable change scores for the Ohio Scales developed by Ogles, Melendez, Davis, & Lunnen (1999) and the clinical significance scores developed by Ogles and Healy (2005) as the criteria. In that study, the lower bound score of 20 on Problem Severity for the Moderate level of care was selected as the clinical cutoff, and the upper bound score of 50 on Functioning for the Moderate level of care was selected as the clinical cutoff.

When both Reliable Change and Clinical Significance conditions are met, the improvement is classified as Reliable and Clinically Significant Change. The Ohio Mental Health Consumer Outcomes System Report 12 includes more information about these concepts and the methods used to compute the Reliable Change and Clinical Significance threshold scores for the measures in the Ohio Consumer Outcomes System. (<http://www.mh.state.oh.us/oper/outcomes/reports/rpt.quarterly.12.pdf>)

Seven categories of changes in this report

Using the above concepts, we developed seven categories of change to sum up changes between two time points. These are Improvement with Clinical Significance, Improvement without Clinical Significance, Partial Improvement, No Change, Partial Deterioration, Deterioration without Clinical Significance, and Deterioration with Clinical Significance.

Improvement with Clinical Significance

Improvement with Clinical Significance refers to those results that show change in a positive direction greater than or equal to the Reliable Change threshold (i.e., 10 for Problem Severity and 8 for Functioning) and at the same time, the records indicate changes from an initial score in the clinical range to a follow-up score in the non-clinical range. For example, an Improvement with Clinical Significance in Problem Severity of the Ohio Scales means there is a reduction of 10 or more points on the scale scores between the two assessments, and at the same time, the scale score is more than the clinical cutoff of Problem Severity (which is 20) at the initial assessment and it is lower than the clinical cutoff of Problem Severity at the follow-up assessment. Simply put, a person experiences a real improvement and moves from the clinical to the non-clinical range.

Improvement without Clinical Significance

Improvement without Clinical Significance here means those results that show changes greater or equal to the Reliable Change (for example, a reduction of 10 in the total Problem Severity scores). However, there is no Clinical Significance. These people may have started in the non-clinical range, or experienced change from an extreme score.

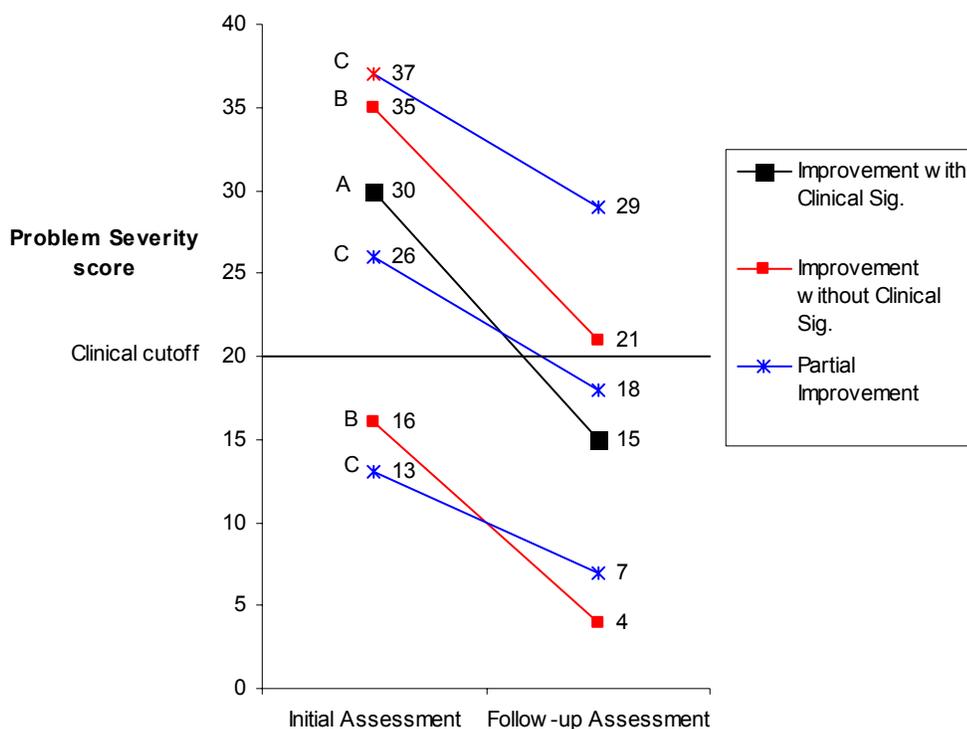
² Jacobson & Truax (1991) and Evans, Margison, & Barkham (1998) provided precise summaries of the calculation with illustrated examples in explaining the two concepts.

Partial Improvement

For Partial Improvement, we take those cases that show a magnitude of improvement between half of the range of Reliable Change (i.e., 5 for Problem Severity and 4 for Functioning) and the full range of Reliable Change (i.e., 10 for Problem Severity and 8 for Functioning). After producing report #12 (see http://www.mh.state.oh.us/oper/outcomes/reports/rpt_quarterly_12.pdf) on Reliable Change and Clinical Significance, we received some comments indicating that the Clinical Significance threshold was too high and that far less change was still important and real. To reflect those comments and recognize where some change has occurred, we have added the partial categories.

The above three groups are illustrated in Figure 1. The Improvement with Clinical Significance (the line labeled A) shows an improvement of 15 and also crossing from the clinical range at initial assessment to the non-clinical range at follow-up assessment. The two Improvements without Clinical Significance are illustrated by the lines labeled B. They show an improvement of 14 and 12 respectively. However, since the two lines are not in the clinical range at initial assessment and in the non-clinical range at follow-up assessment, they are not considered as Improvement with Clinical Significance. The three partial improvements (with lines labeled C) show a reduction of 8, 8 and 6 respectively. Even though one of them has crossed the clinical cutoff point between the two assessments, the magnitude of the change does not achieve the Reliable Change criterion and so it is only considered as Partial Improvement.

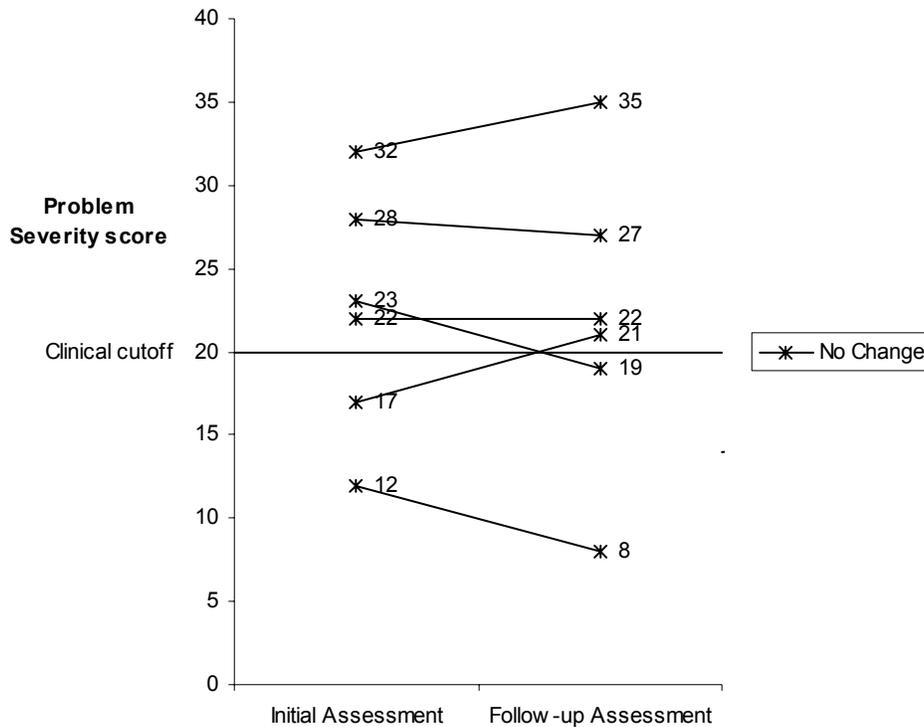
Figure 1. Illustration for Improvement Categories



No Change

The No Change group is the group that shows improvement or deterioration of less than half of the Reliable Change threshold. Figure 2 illustrates some possible scenarios of this group. All lines in Figure 2 show an improvement or deterioration of less than half of the Reliable Change (i.e., less than 5 in Problem Severity score). Whether they are in the clinical range or the non-clinical range, or whether there are any changes from clinical to non-clinical or vice versa will not affect the classification.

Figure 2. Illustration for No Change Categories



Partial Deterioration

Similar to Partial Improvement, Partial Deterioration is a group of cases with deterioration of more than half the range of Reliable Change (i.e., 5 in Problem Severity score) but less than a full Reliable Change (i.e., 10 in Problem Severity score) in the deterioration direction. The lines labeled D in Figure 3 are showing Partial Deterioration, with deterioration of 7, 6, and 6 respectively. Even though one of them has crossed the clinical cutoff point between the two assessments, the magnitude of the change does not achieve the Reliable Change criterion and so it is only considered as Partial Deterioration.

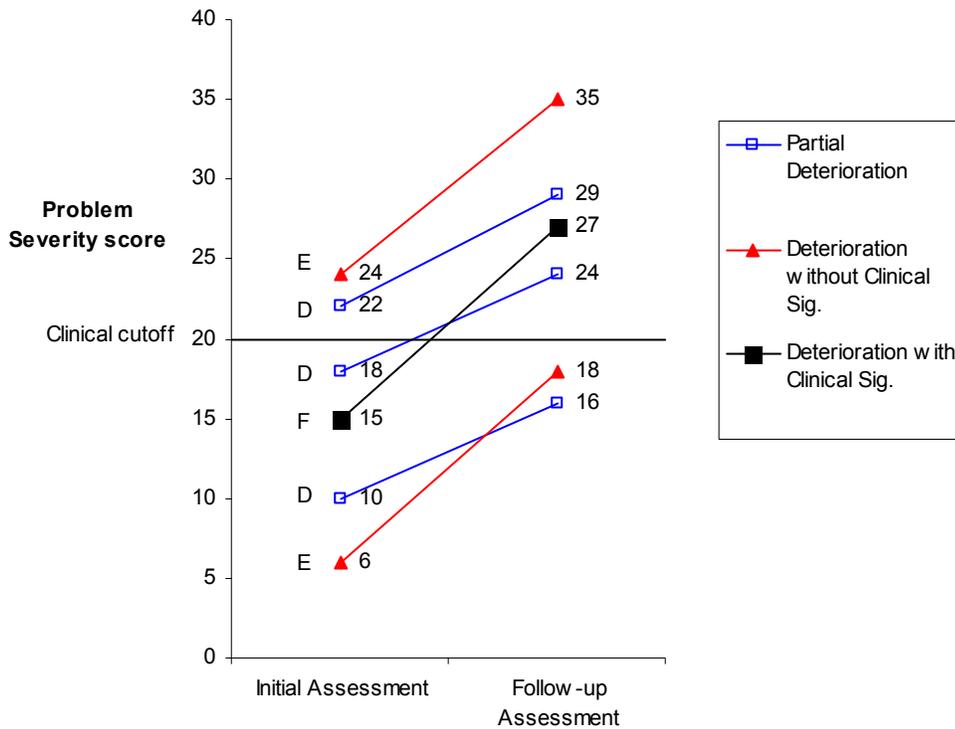
Deterioration without Clinical Significance

The Deterioration without Clinical Significance is the group with a deterioration greater than the Reliable Change (e.g. an increase of 10 in Problem Severity score, which is a deterioration), but they do not have both an initial score in the non-clinical range and a follow-up score in the clinical range. The Deterioration without Clinical Significance is illustrated by the lines labeled E in Figure 3.

Deterioration with Clinical Significance

A Deterioration with Clinical Significance is the group with deterioration between the two assessments greater than the Reliable Change threshold and at the same time showing an initial score in the non-clinical range and a follow-up score in the clinical range. The line labeled F in Figure 3 shows a Deterioration with Clinical Significance with an increase of 12 in the Problem Severity score and changing from the non-clinical range at initial assessment to the clinical range at follow-up assessment.

Figure 3. Illustration for Deterioration Categories



For those who would like to compare local results to the benchmarks reports here, but only have three categories (Reliable Improvement, No Change, Reliable Deterioration), you can add together the percentages of the Improvement with Clinical Significance and Improvement without Clinical Significance categories to come up with the equivalent of the Reliable Improvement category. Similarly, you can add together the percentages of the Deterioration with Clinical Significance and Deterioration without Clinical Significance to come up with the equivalent of the Reliable Deterioration category. Finally, the No Change category can be computed by adding the percentages of the Partial Improvement, No Change and Partial Deterioration categories.

The present report employed the August 10, 2007 Production Data in the Ohio Mental Health Consumer Outcomes System. In the following analyses, four pair-wise matching datasets (Initial and 3-month assessments, Initial and 6-month assessments, Initial and 1-year assessments, and Initial and Termination assessments)³ are constructed from all unique individuals in the system with valid Outcomes measures (i.e., Problem Severity or Functioning) in both the Initial assessment and a follow-up assessment (i.e., 3-month assessment, 6-month assessment, 1-year assessment, and the assessment marked as termination respectively).

³ The Initial assessment period includes all assessments administered within 44 days from the date of admission to service. The 3-month assessment includes all assessments administered between 45 days and 135 days. The 6-month period includes all assessments administered within 136 to 227 days from the date of admission. The one-year period includes all assessments administered within 320 days to 410 days from the date of admission to service. And the termination assessment includes valid Outcomes records marked as termination administration. If more than one assessment was completed within a time period, the earlier valid assessment would be used for that time period.

Multivariate Analysis: Exploring factors predicting changes

In assessing the contribution of various factors and covariates to changes in Outcomes, we used the magnitude of changes between the Initial assessment and the follow-up assessment in Problem Severity and Functioning of the Ohio Scales respectively as the dependent variable in a series of multivariate analyses. The following factors were included: intake score, age, county size, race, gender, and primary diagnostic group, as well as all interaction terms. Results showed that county size, intake Outcomes level (Problem Severity and Functioning respectively), and age are significant factors, with all interaction effects taken into consideration. Appendix A contains tables reporting significant main effects and interaction terms affecting the magnitude of change in the Problem Severity and Functioning.

To be consistent with the other benchmark reports, breakouts for age group and gender are included regardless of whether they are significant factors. In addition, breakouts for all significant ($p < .05$) main effects are included.

Discussion

Ogles, Lunnen & Bonesteel (2001) cite a variety of definitions used in research on clinically significant change. In this report, we have chosen a narrow definition of clinically significant change, where the client moves from one distribution to another with a magnitude of change above the threshold of the reliable change index (Jacobson, Follette, & Revenstorf, 1984, 1986). Thus, in this report, those clients who achieved a reliable improvement but started in the non-clinical range (and therefore did not cross the cutoff point) are classified in the "Improvement without Clinical Significance" category rather than the "Improvement with Clinical Significance" category. Our decision preserves clinical significance as a movement across the clinical cutting score, rather than a classification based on the status of being in the non-clinical range at the end of treatment. This decision has practical implications for the appearance of the results: about 3% of the youth included in this report started in the non-clinical range on Problem Severity and made a reliable improvement, and at the other end of the continuum, about 5% of youth started in the non-clinical range and experienced a reliable deterioration. We think this classification system makes a more conservative statement about treatment effectiveness, and will be more widely understood by researchers using clinical significance.

This report presents data in a new way, but does not present new data. This view of change over time complements the views in the odd-numbered reports in our Quarterly Report series (<http://www.mh.state.oh.us/oper/outcomes/reports.quarterly.html>) and available in the Outcomes Data Mart (available through: <http://www.mh.state.oh.us/oper/outcomes/data.mart.index.html>). We believe this new view of data gives the user many points of comparison for use in benchmarking aggregate data. Several points are important to make as we begin to understand these data. First, the data reported here represent a subset of all of the cases that began treatment. The majority of youth who begin treatment do not make it to the three-month administration time period. Further, for those who do make it to the 3-month time period, not all Outcomes data are reported. Second, not all of the factors that may shape Outcomes change scores were included in this analysis, only those factors that were included in the Outcomes data.

More generally, the data belie the saying that "one third get better, on third get worse, and one third stay the same." Clearly, when considering all of the improvement categories, almost one half get better. This is welcome news about the state of treatment, but *cannot* be taken as a stamp of approval for the status quo. Many people dropped out before they could be included in this assessment, and too many still do not make progress or actually get worse in treatment. Every indication is that more can be done to increase the percentage of youth who get better, by improving the methods for engaging consumers, including engagement around their Outcomes data, the use of Evidence-Based Practices, and the use of quality improvement efforts. Wampold's (2001) research indicates that only about 15% of the variance in the outcomes of psychotherapy can be accounted for by the modality of treatment. However, that 15%, can be maximized by implementing the best-evidenced practices that are available. Wampold also indicates that fully 60% of the variance in psycho-therapy outcomes can be explained by consumer factors, such as readiness to change. Much can be done to understand where a consumer is and adapt treatment

to fit the consumer's desires and readiness to change. Harmon, Lambert, and Smart (2007) show how the rate of deteriorating cases can be decreased using outcomes feedback. Further, quality assurance and quality improvement efforts can go a long way to assure that clinicians maintain treatment fidelity to standards, and that provider agencies are obtaining feedback about effectiveness and looking for ways to improve.

This report represents baseline measures for the system. What future results Ohio's public mental health system obtains will depend on the ability of the system to improve treatment methods and factors like consumer engagement.

Understanding the charts

Each chart shows the percentage of consumers who fall into one of the seven categories described above: 1) Improvement with Clinical Significance, 2) Improvement Without Clinical Significance, 3) Partial Improvement, 4) No Change, 5) Partial Deterioration, 6) Deterioration without Clinical Significance and Deterioration with Clinical Significance. The percentages shown add to 100% with some differences due to rounding. Some of the charts also break the information out by various demographic variables. Several pieces of information are included with each chart. The number of cases, or N (shown in parentheses behind the title), shows the number of cases that are included in the chart; multiplying the percentage of a category times the N will give the number of cases in that category. For example, the first chart includes an N of 10,111 and 21.8% in the Improvement with Clinical Significance category. Multiplying these numbers ($10,111 \times .218$) shows that about 2,204 youth of the 10,111 youth included experienced Improvement with Clinical Significance. A note about the average number of days between the assessments is included so the reader can see the average days elapsed between assessments. The standard deviation of the days between assessments, which can be understood as the mean distance from the average, is noted in parentheses as the S.D. The standard deviation shows that not all of the data were collected exactly on schedule.

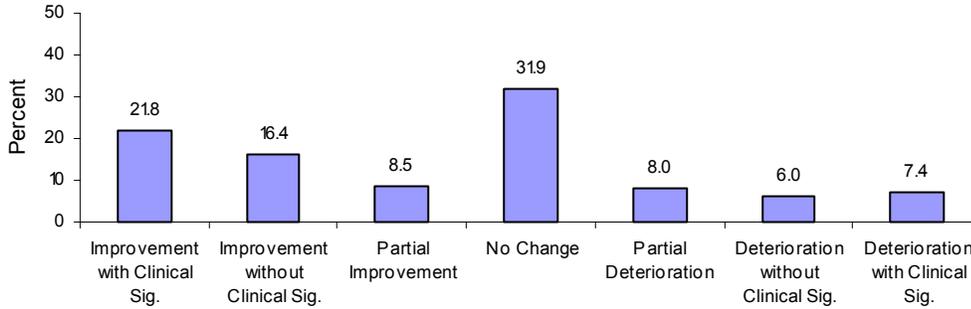
Any questions about the report can be directed to either Kwok Kwan Tam at tamk@mh.state.oh.us, 614-752-9706, or to Jim Healy, at healyj@mh.state.oh.us, 614-752-9311.

Change in Ohio Scales – Problem Severity

Problem Severity—Overall Changes

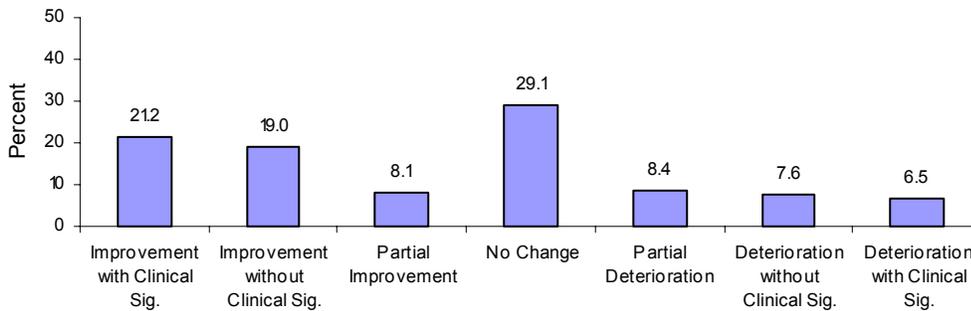
Overall change in Problem Severity, Initial Assessment to 3-month Assessment

Youth Rating - Problem Severity (Initial - 3-month) (N=10,111)



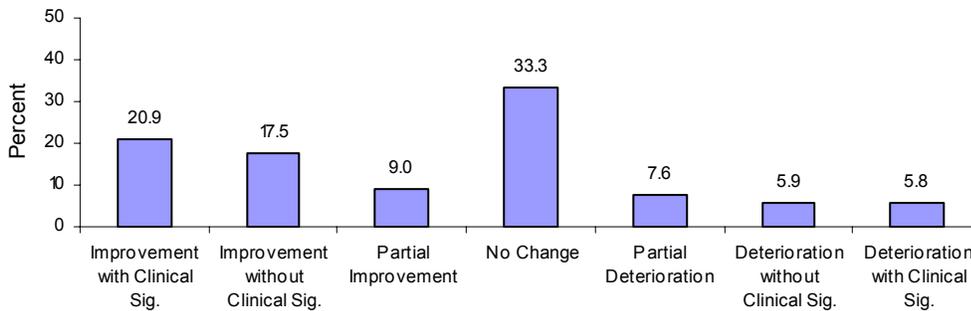
Note: Average number of days between assessments = 94 (S.D.=23)

Parent Rating - Problem Severity (Initial - 3-month) (N=18,531)



Note: Average number of days between assessments = 94 (S.D.=23)

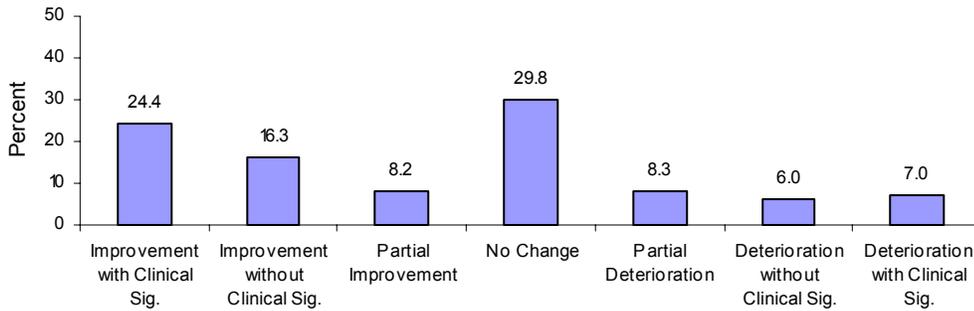
Worker Rating - Problem Severity (Initial - 3-month) (N=29,580)



Note: Average number of days between assessments = 93 (S.D.=23)

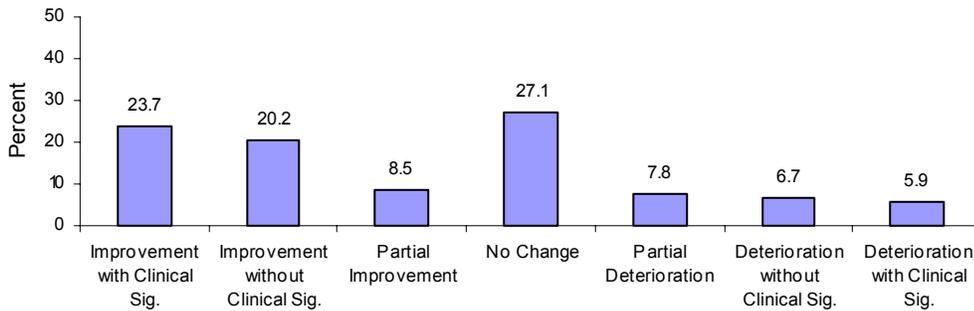
Overall Change in Problem Severity, Initial Assessment to 6-month Assessment

Youth Rating - Problem Severity (Initial - 6-month) (N=14,723)



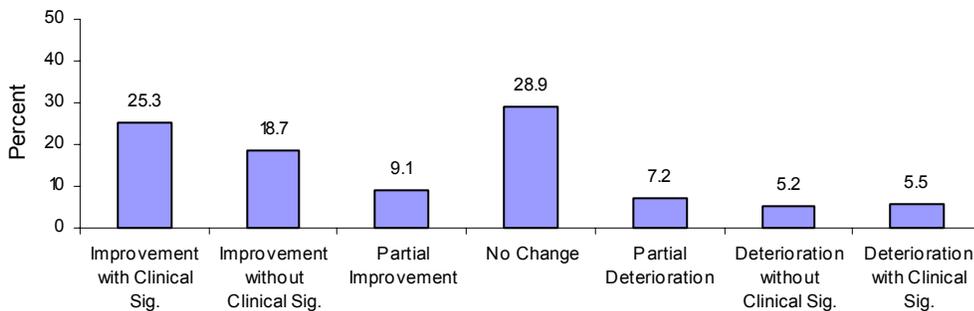
Note: Average number of days between assessments = 194 (S.D.=32)

Parent Rating - Problem Severity (Initial - 6-month) (N=28,541)



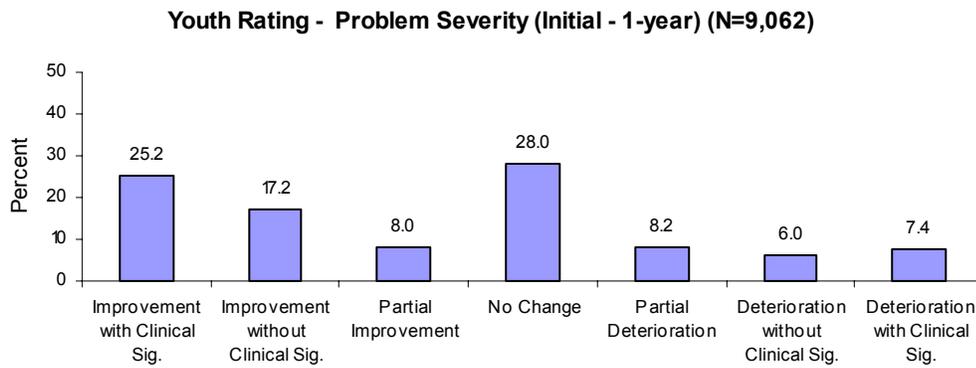
Note: Average number of days between assessments = 195 (S.D.=32)

Worker Rating - Problem Severity (Initial - 6-month) (N=40,635)

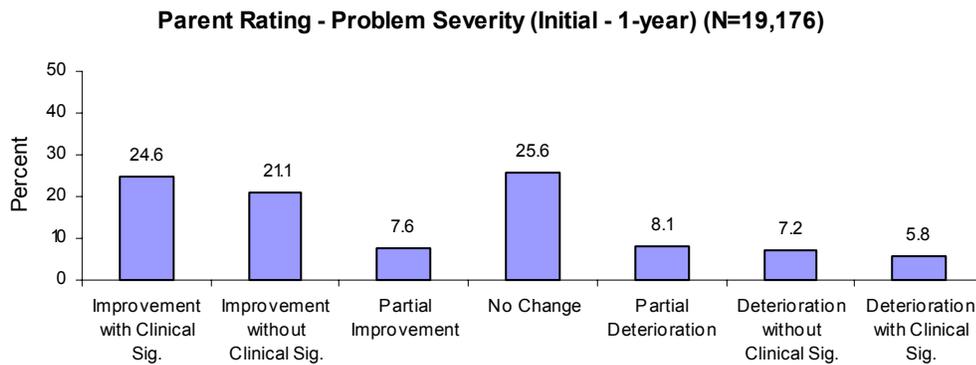


Note: Average number of days between assessments = 193 (S.D.=33)

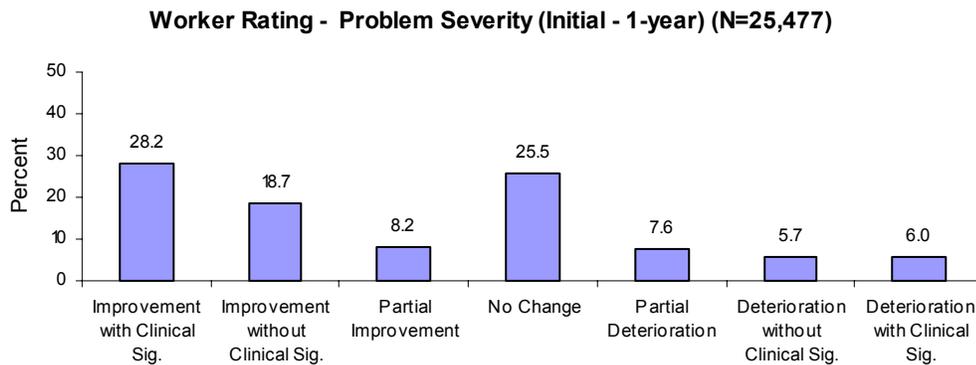
Overall change in Problem Severity, Initial Assessment to 1-year Assessment



Note: Average number of days between assessments = 356 (S.D.=46)



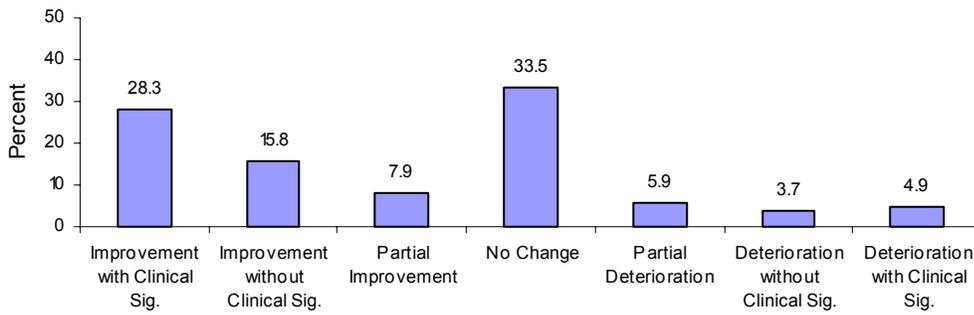
Note: Average number of days between assessments = 356 (S.D.=46)



Note: Average number of days between assessments = 352 (S.D.=46)

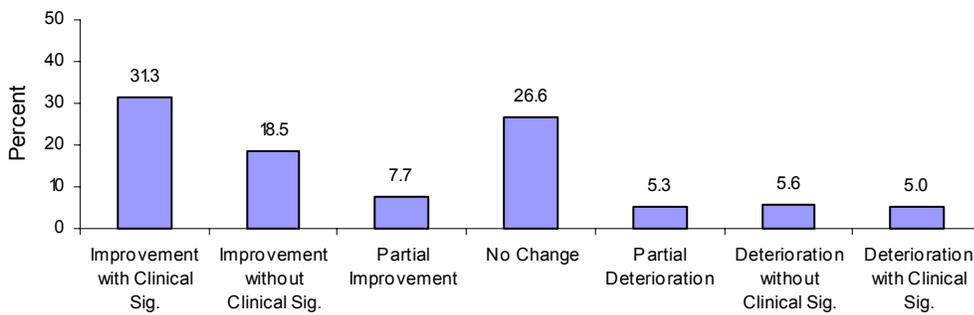
Overall change in Problem Severity, Initial Assessment to Termination

Youth Rating - Problem Severity (Initial - Termination) (N=8,012)



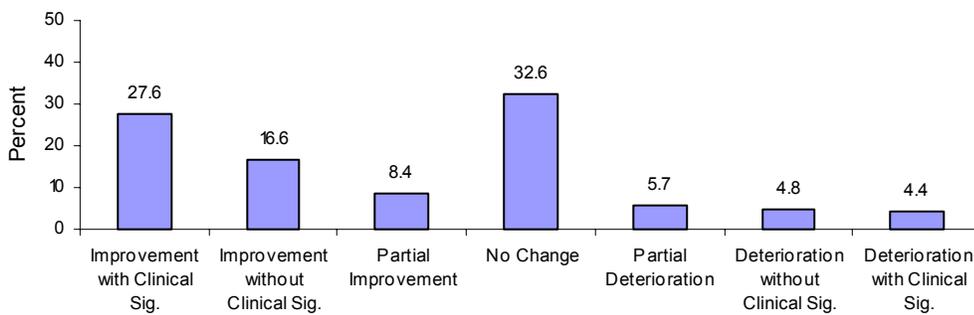
Note: Average number of days between assessments = 305 (S.D.=395)

Parent Rating - Problem Severity (Initial - Termination) (N=13,150)



Note: Average number of days between assessments = 332 (S.D.=381)

Worker Rating - Problem Severity (Initial - Termination) (N=42,288)

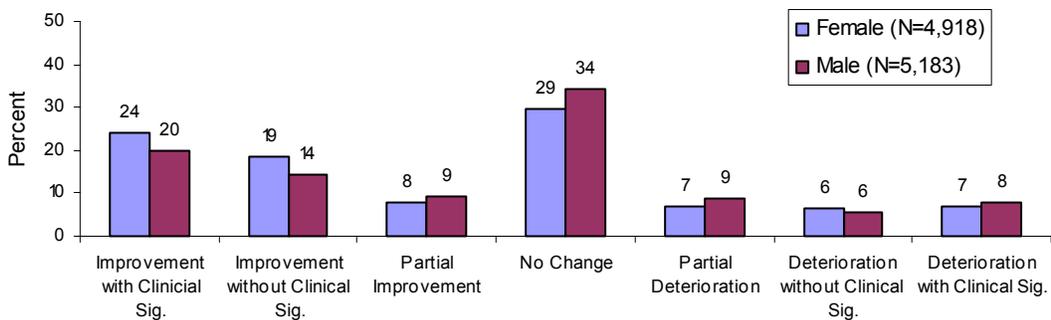


Note: Average number of days between assessments = 300 (S.D.=345)

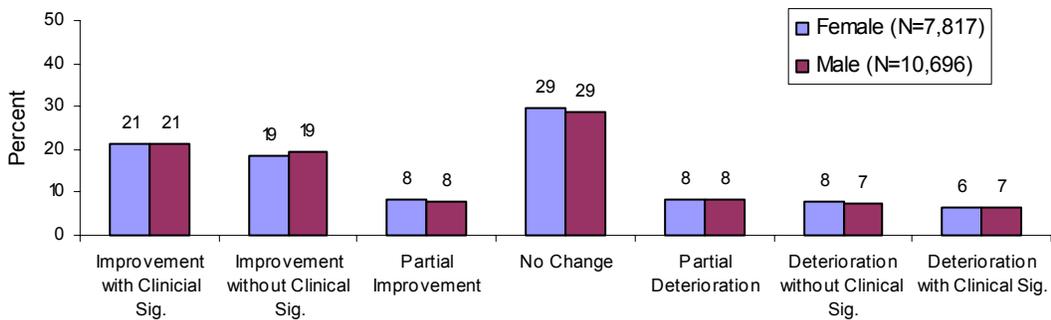
Changes by Gender—Problem Severity

Change in Problem Severity by Gender, Initial Assessment to 3-month Assessment

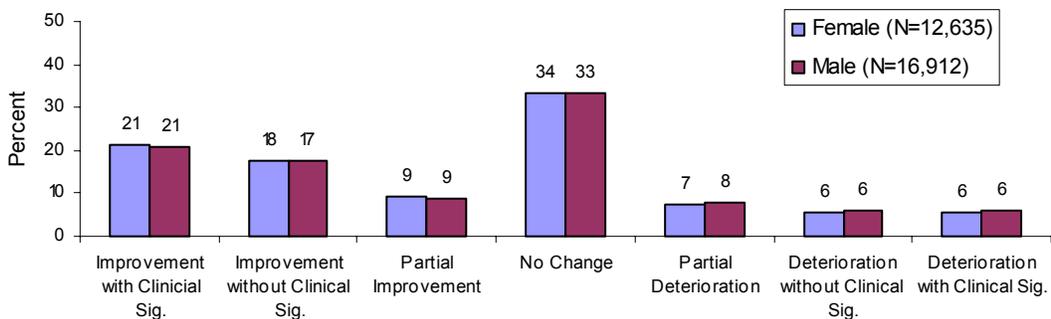
Youth Rating - Problem Severity (Initial - 3-month) by Gender



Parent Rating - Problem Severity (Initial - 3-month) by Gender

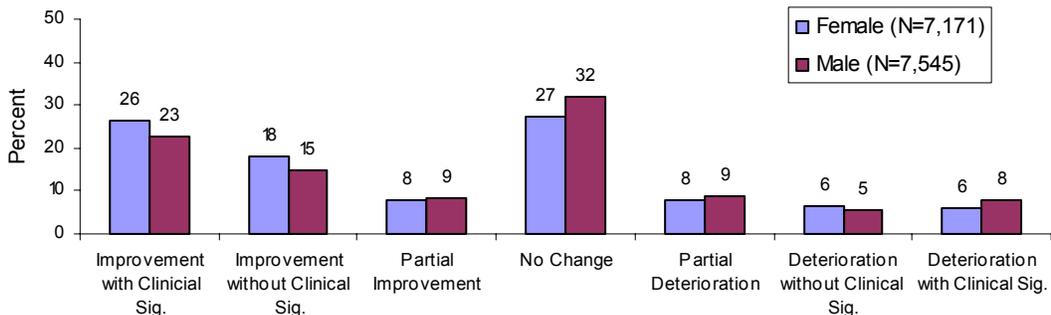


Worker Rating - Problem Severity (Initial - 3-month) by Gender

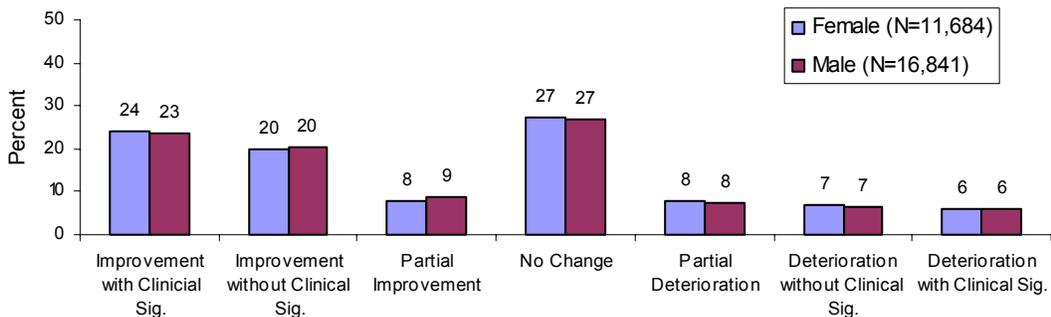


Change in Problem Severity by Gender, Initial Assessment to 6-month Assessment

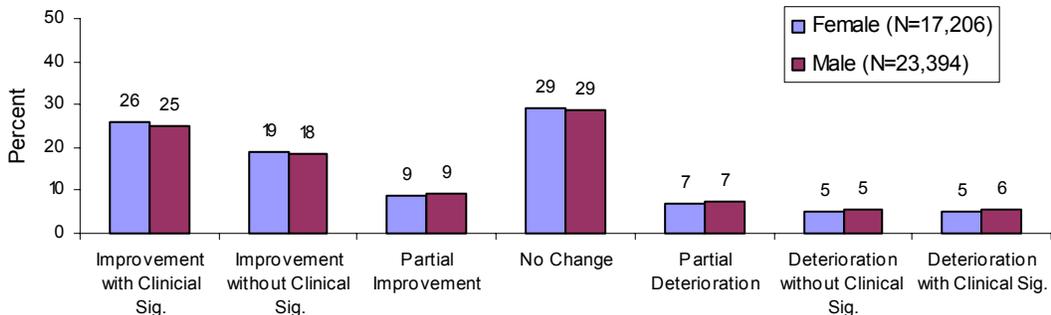
Youth Rating - Problem Severity (Initial - 6-month) by Gender



Parent Rating - Problem Severity (Initial - 6-month) by Gender

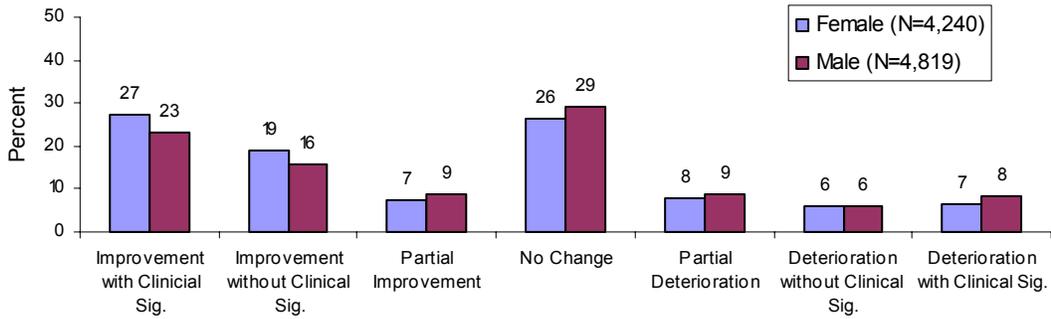


Worker Rating - Problem Severity (Initial - 6-month) by Gender

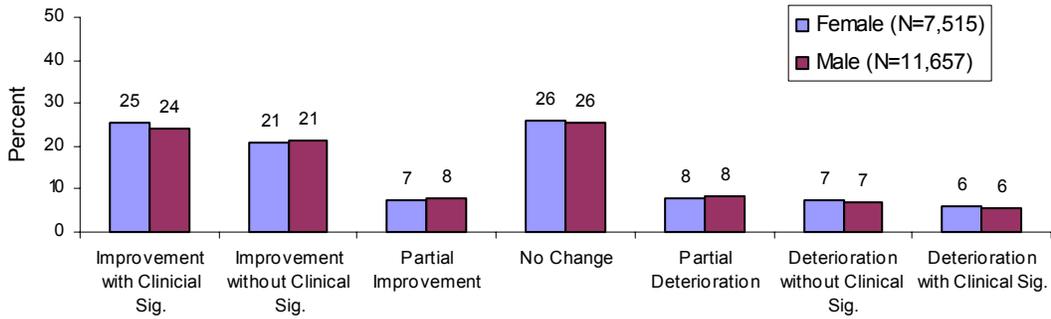


Change in Problem Severity by Gender, Initial Assessment to 1-year Assessment

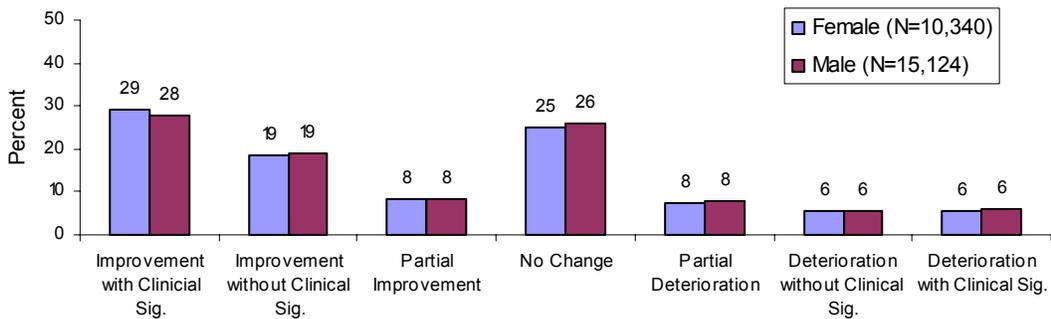
Youth Rating - Problem Severity (Initial - 1-year) by Gender



Parent Rating - Problem Severity (Initial - 1-year) by Gender

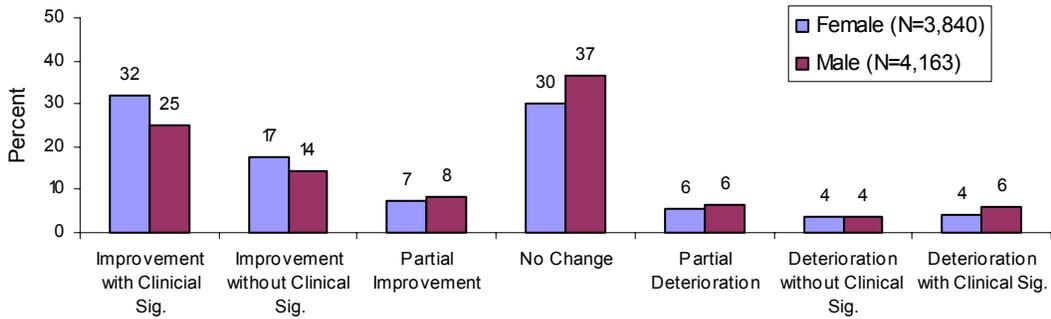


Worker Rating - Problem Severity (Initial - 1-year) by Gender

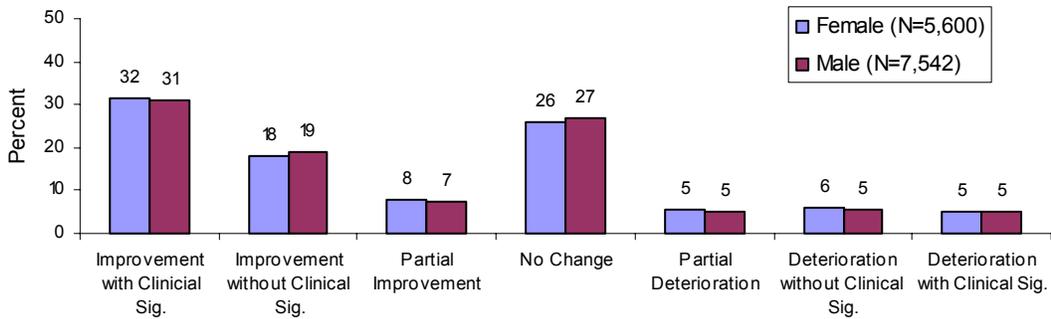


Change in Problem Severity by Gender, Initial Assessment to Termination

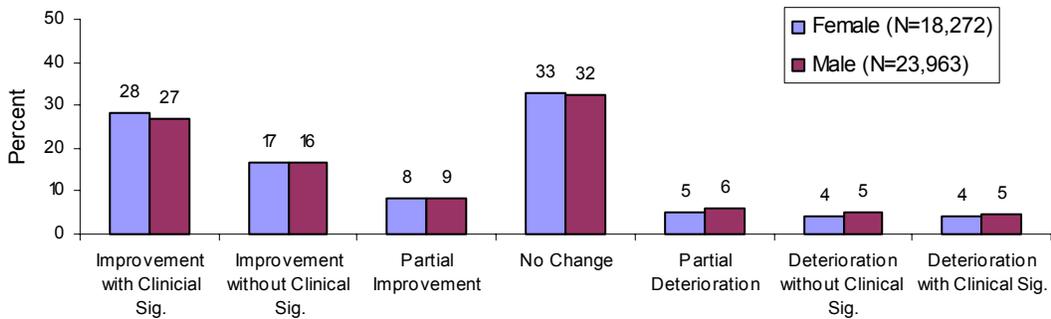
Youth Rating - Problem Severity (Initial - Termination) by Gender



Parent Rating - Problem Severity (Initial - Termination) by Gender



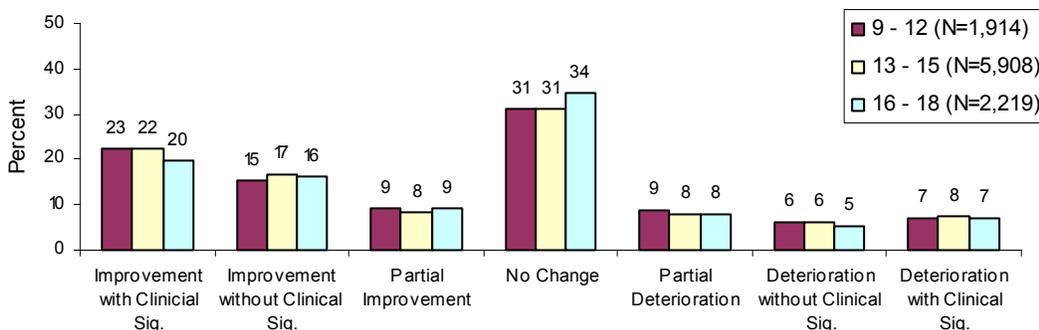
Worker Rating - Problem Severity (Initial - Termination) by Gender



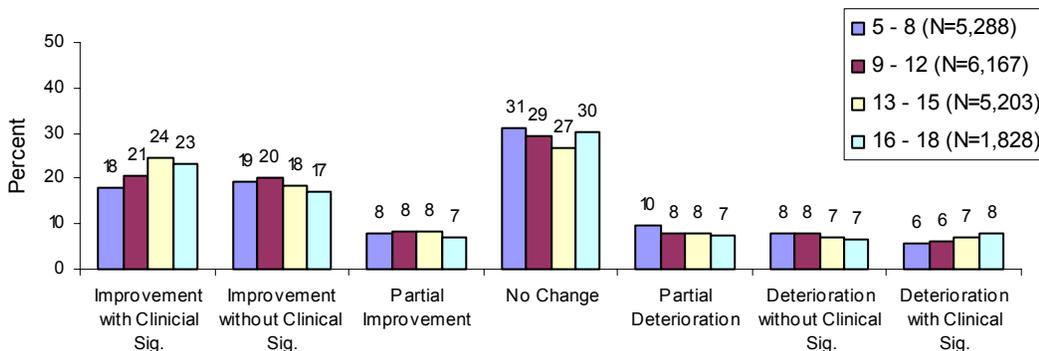
Changes by Age—Problem Severity

Change in Problem Severity by Age, Initial Assessment to 3-month Assessment

Youth Rating - Problem Severity (Initial - 3-month) by Age

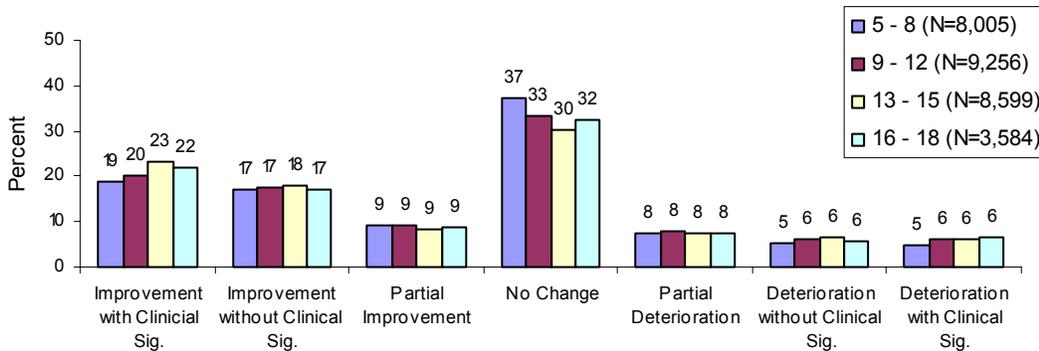


Parent Rating - Problem Severity (Initial - 3-month) by Age



Note: Age is a significant covariate in multivariate analysis with Problem Severity as dependent variable.

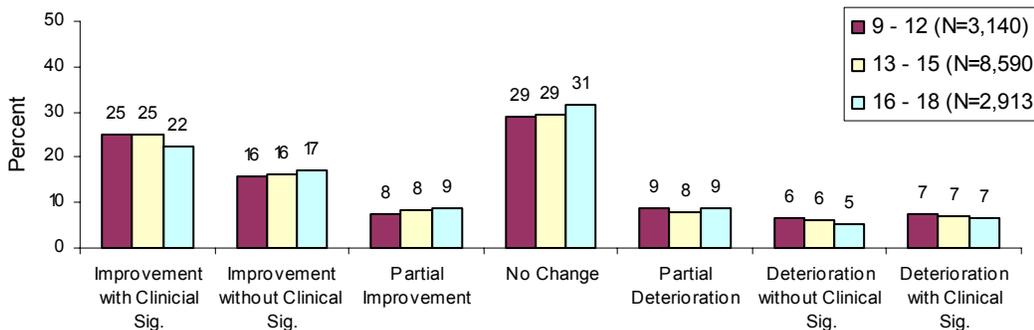
Worker Rating - Problem Severity (Initial - 3-month) by Age



Note: Age is a significant covariate in multivariate analysis with Problem Severity as dependent variable.

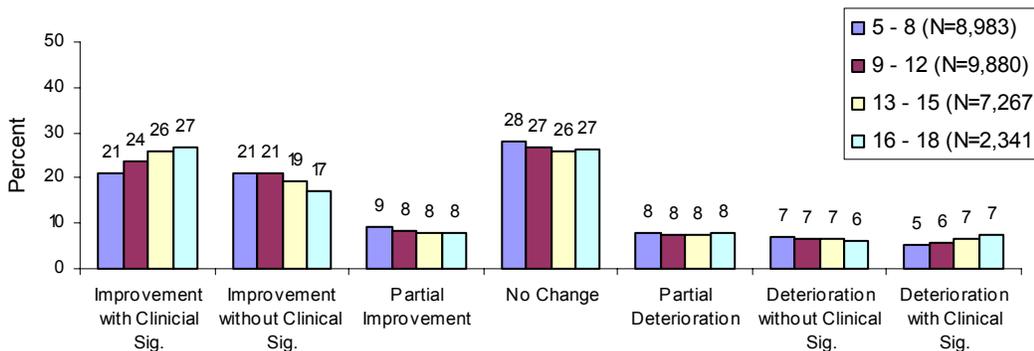
Change in Problem Severity by Age, Initial Assessment to 6-month Assessment

Youth Rating - Problem Severity (Initial - 6-month) by Age



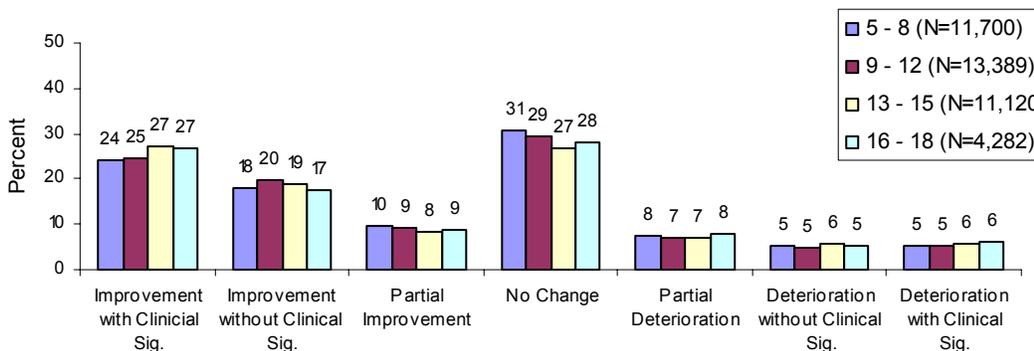
Note: Age is a significant covariate in multivariate analysis with Problem Severity as dependent variable.

Parent Rating - Problem Severity (Initial - 6-month) by Age



Note: Age is a significant covariate in multivariate analysis with Problem Severity as dependent variable.

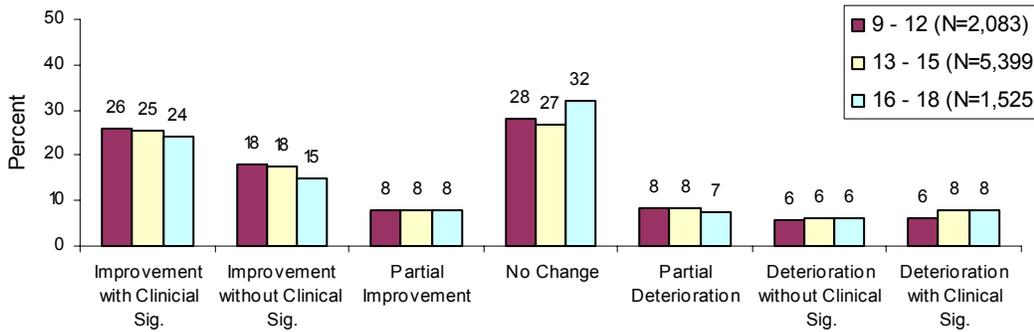
Worker Rating - Problem Severity (Initial - 6-month) by Age



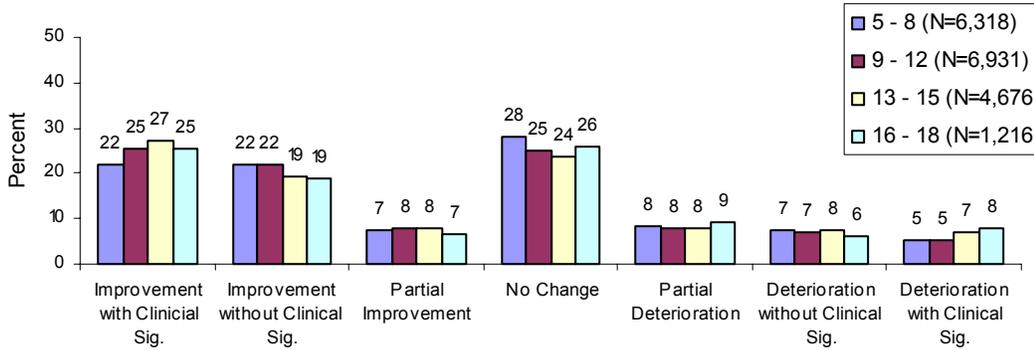
Note: Age is a significant covariate in multivariate analysis with Problem Severity as dependent variable.

Change in Problem Severity by Age, Initial Assessment to 1-year Assessment

Youth Rating - Problem Severity (Initial - 1-year) by Age

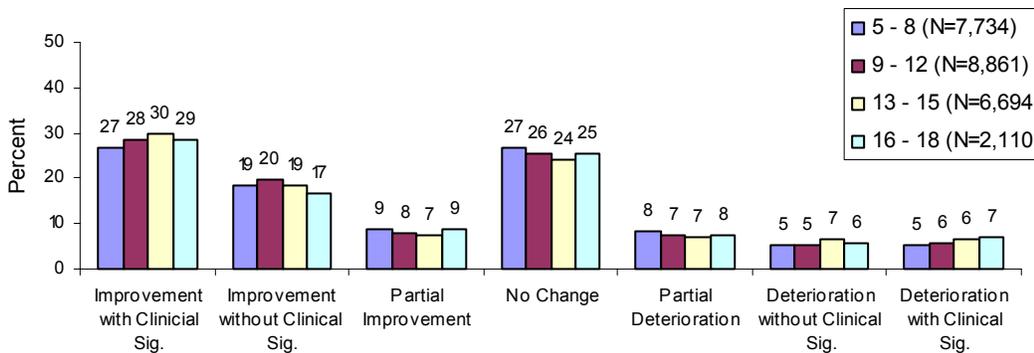


Parent Rating - Problem Severity (Initial - 1-year) by Age



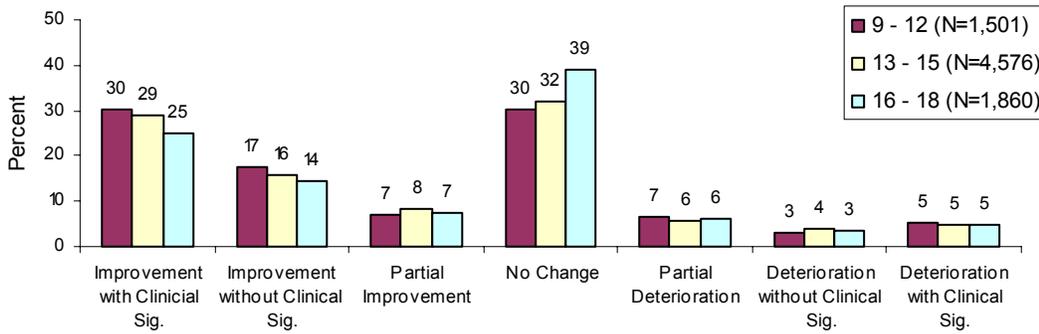
Note: Age is a significant covariate in multivariate analysis with Problem Severity as dependent variable.

Worker Rating - Problem Severity (Initial - 1-year) by Age

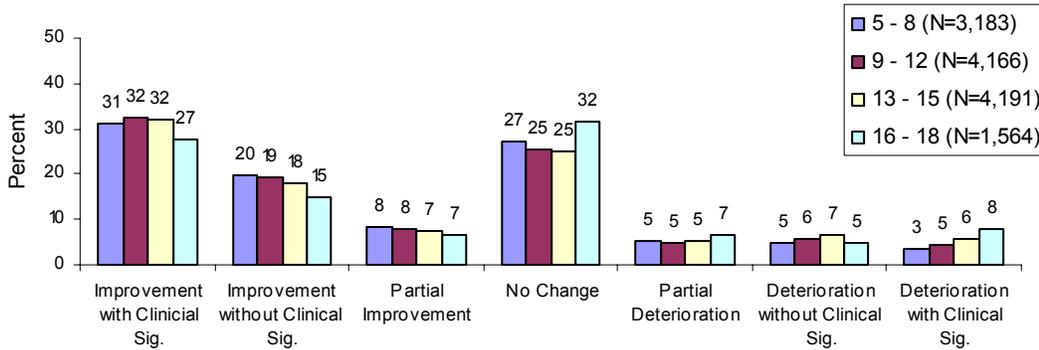


Change in Problem Severity by Age, Initial Assessment to Termination

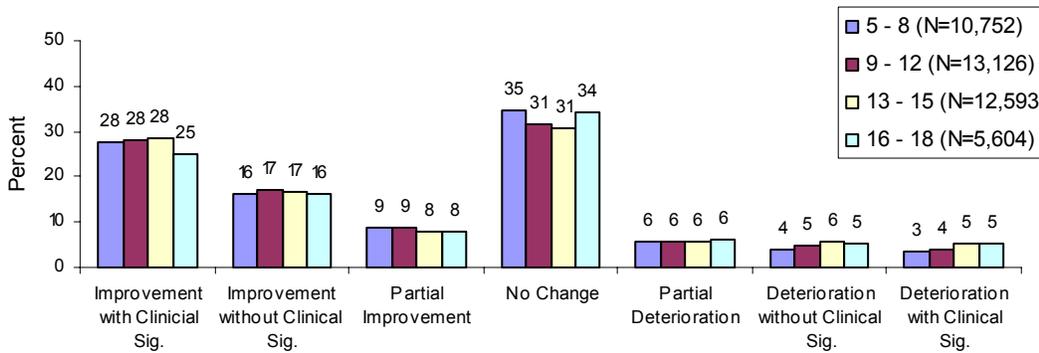
Youth Rating - Problem Severity (Initial - Termination) by Age



Parent Rating - Problem Severity (Initial - Termination) by Age



Worker Rating - Problem Severity (Initial - Termination) by Age

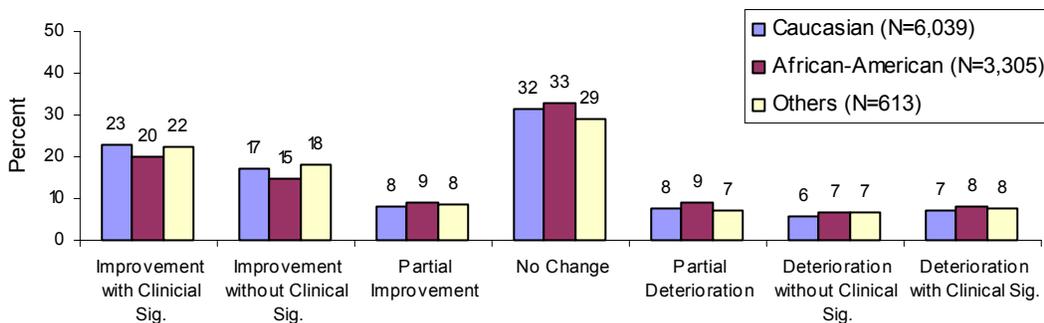


Note: Age is a significant covariate in multivariate analysis with Problem Severity as dependent variable.

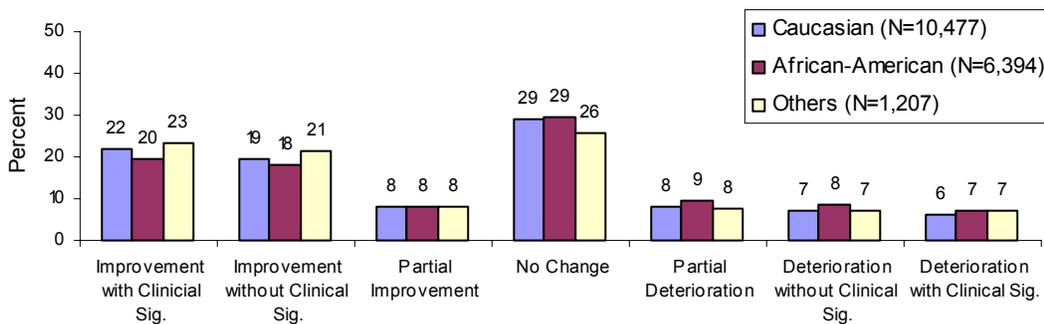
Changes by Race—Problem Severity

Change in Problem Severity by Race, Initial Assessment to 3-month Assessment

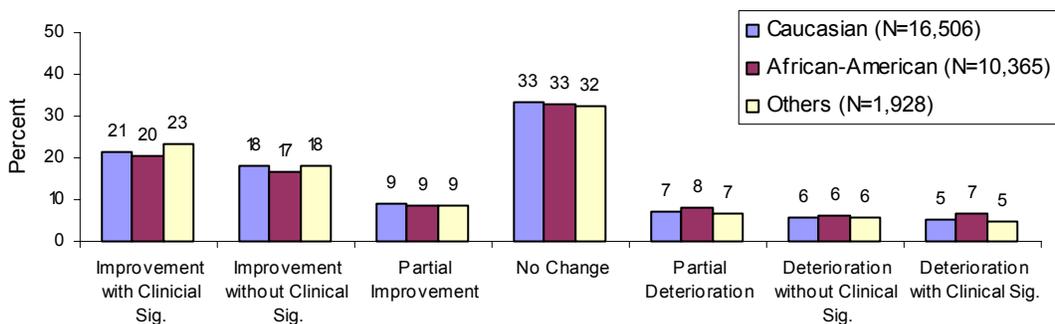
Youth Rating - Problem Severity (Initial - 3-month) by Race



Parent Rating - Problem Severity (Initial - 3-month) by Race

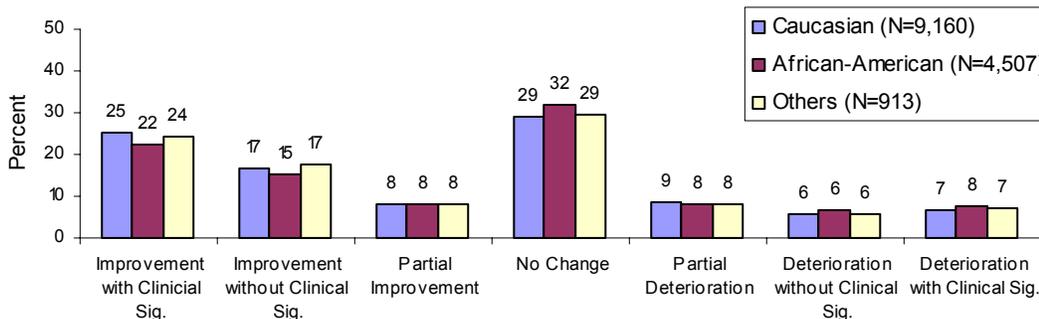


Worker Rating - Problem Severity (Initial - 3-month) by Race

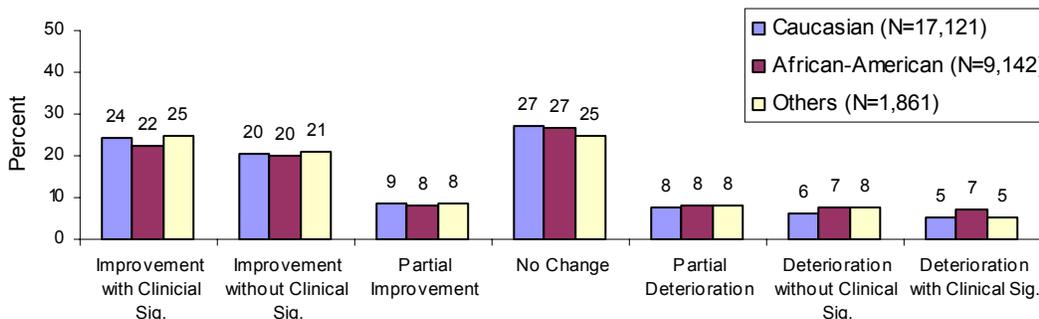


Change in Problem Severity by Race, Initial Assessment to 6-month Assessment

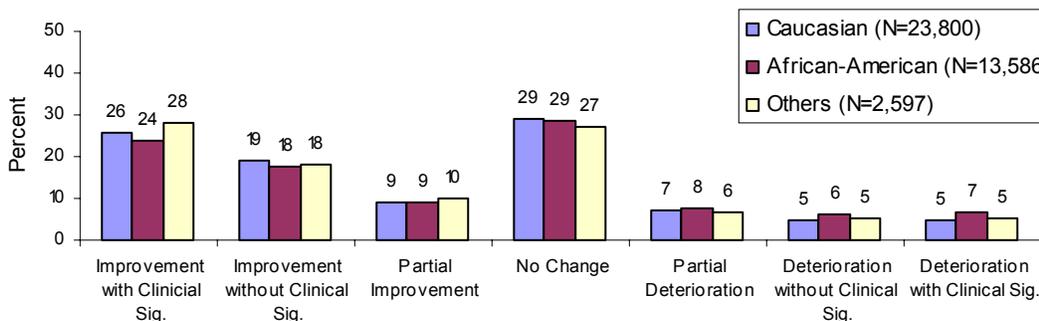
Youth Rating - Problem Severity (Initial - 6-month) by Race



Parent Rating - Problem Severity (Initial - 6-month) by Race

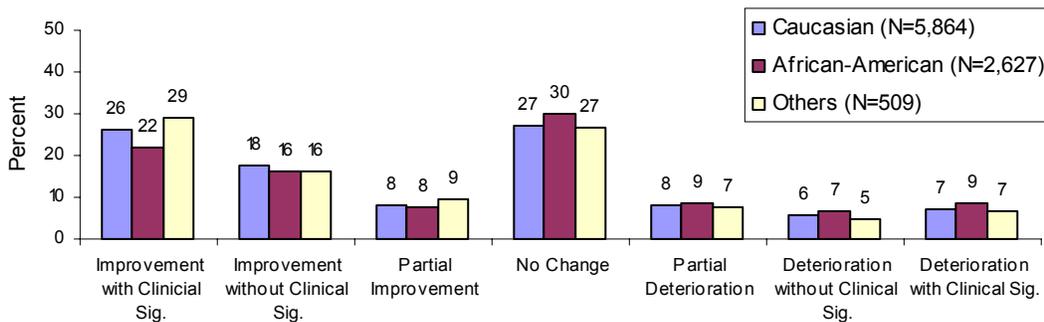


Worker Rating - Problem Severity (Initial - 6-month) by Race

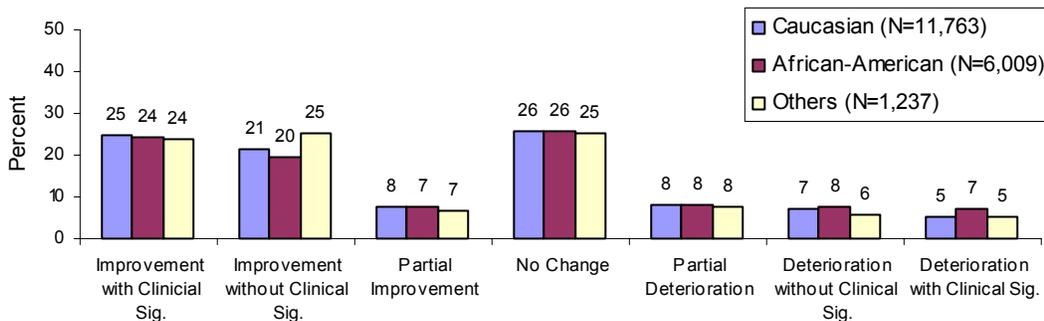


Change in Problem Severity by Race, Initial Assessment to 1-year Assessment

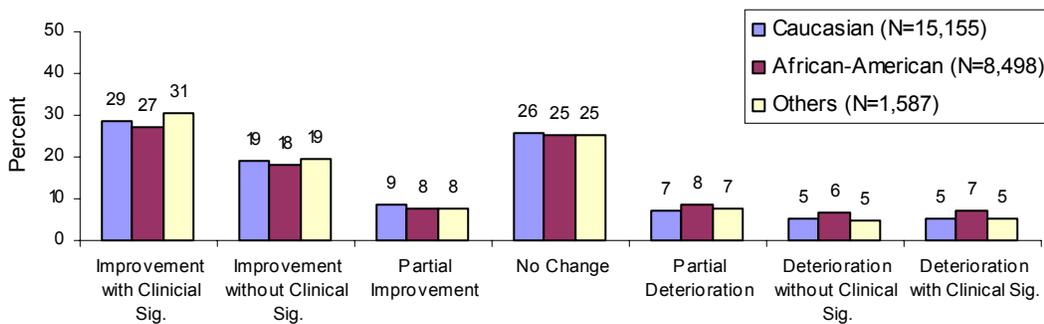
Youth Rating - Problem Severity (Initial - 1-year) by Race



Parent Rating - Problem Severity (Initial - 1-year) by Race

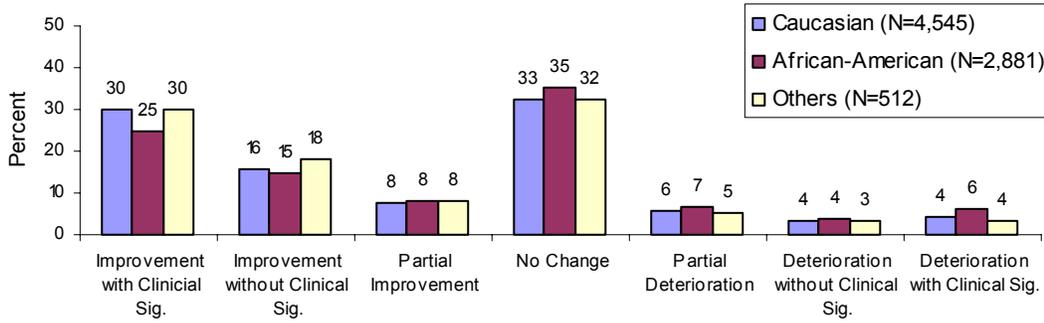


Worker Rating - Problem Severity (Initial - 1-year) by Race

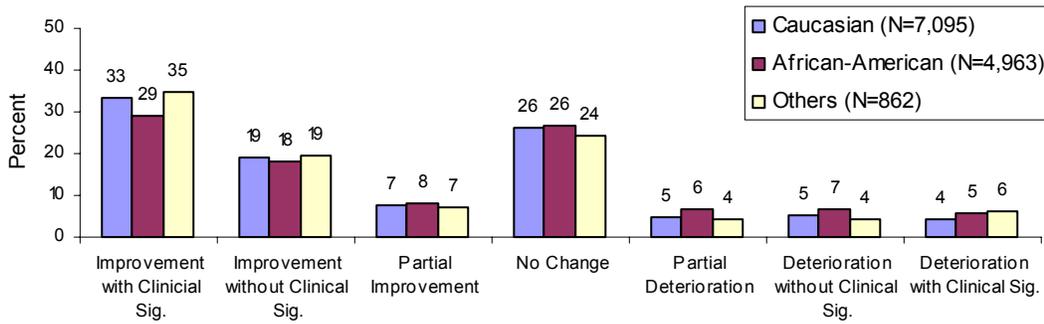


Change in Problem Severity by Race, Initial Assessment to Termination

Youth Rating - Problem Severity (Initial - Termination) by Race

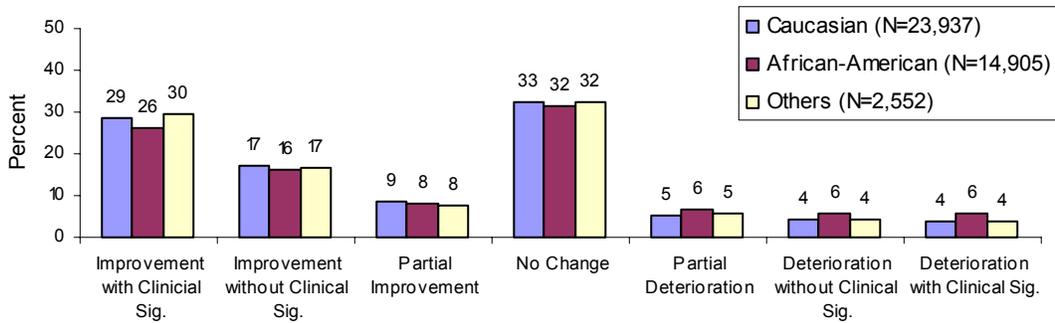


Parent Rating - Problem Severity (Initial - Termination) by Race



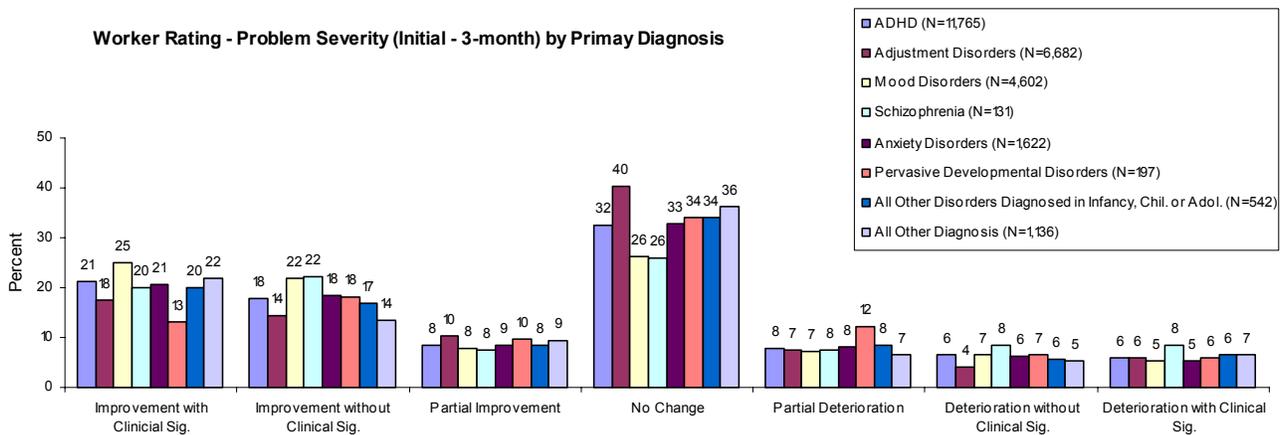
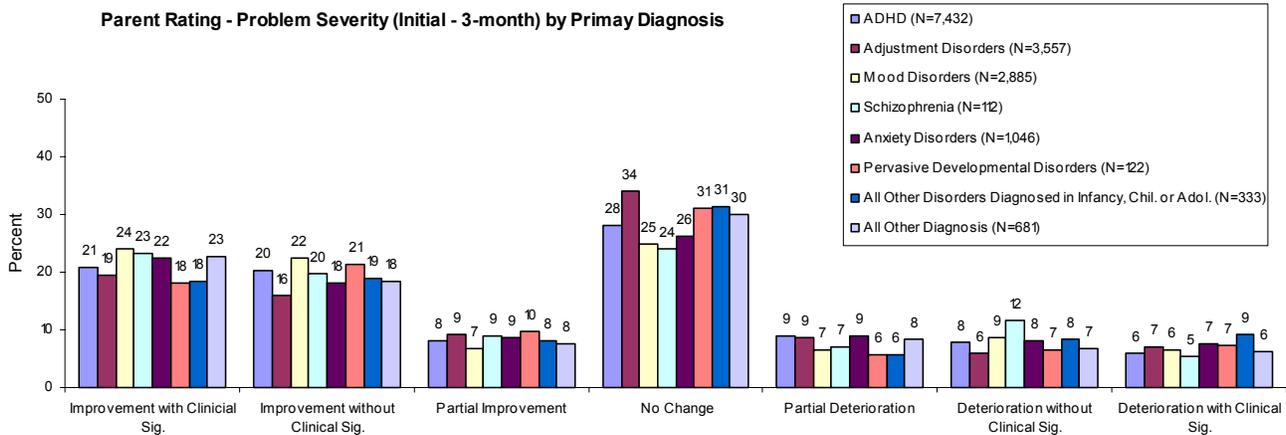
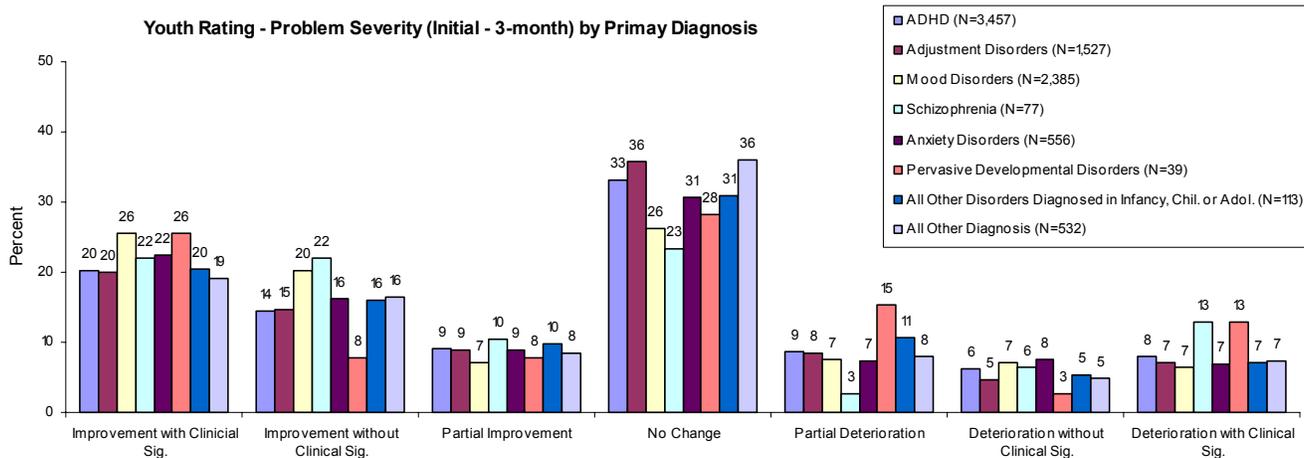
Note: Race is a significant factor in multivariate analysis with Problem Severity as dependent variable.

Worker Rating - Problem Severity (Initial - Termination) by Race

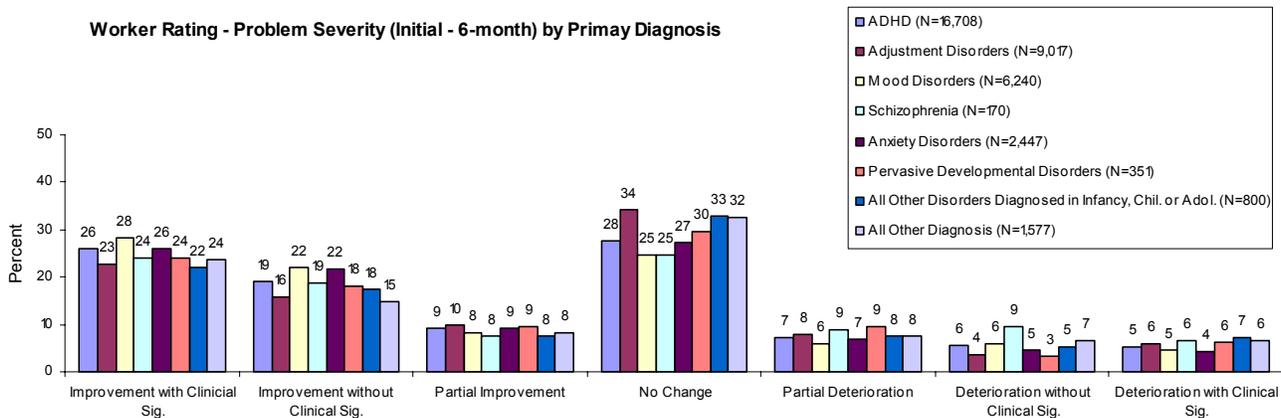
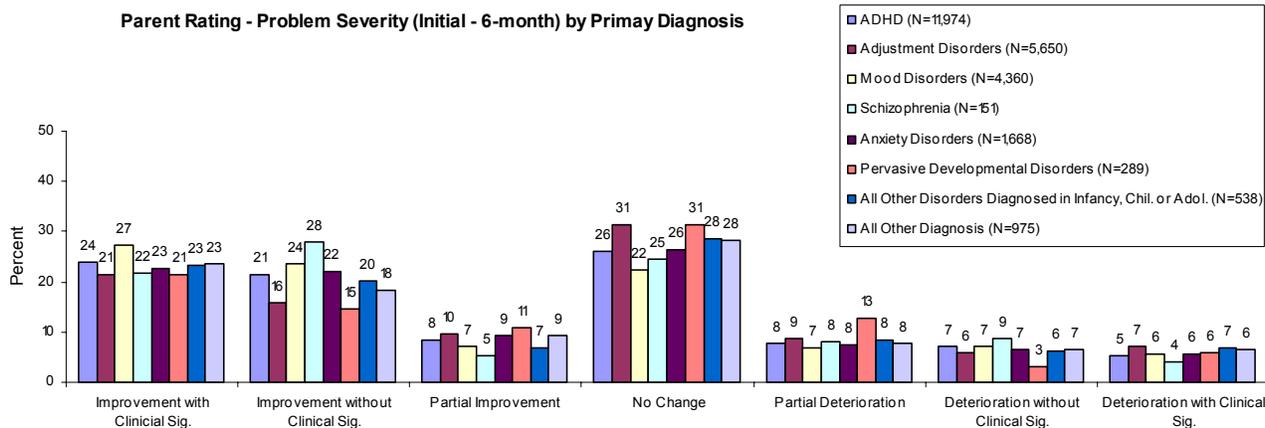
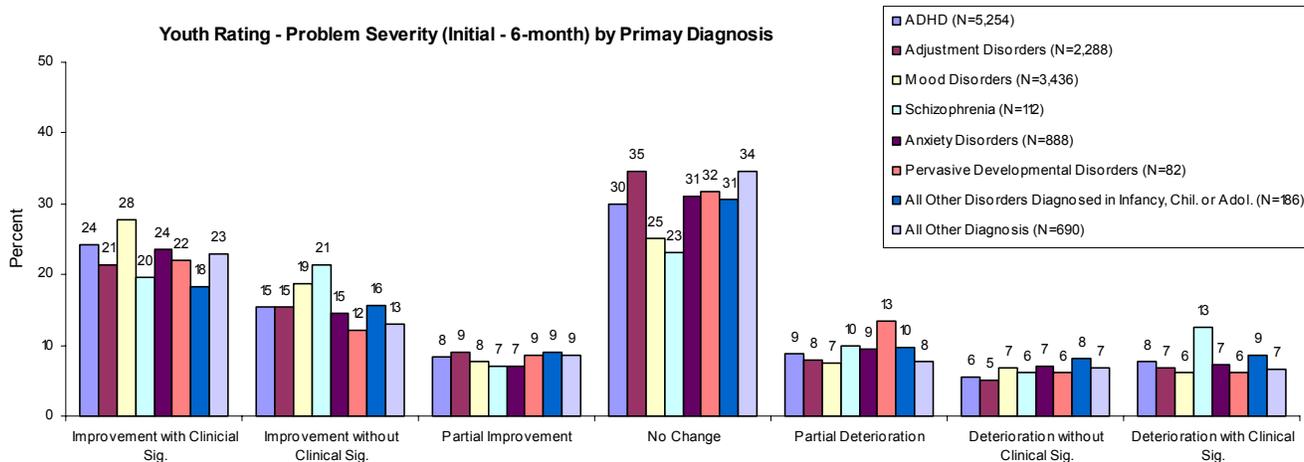


Changes by Diagnostic Group—Problem Severity

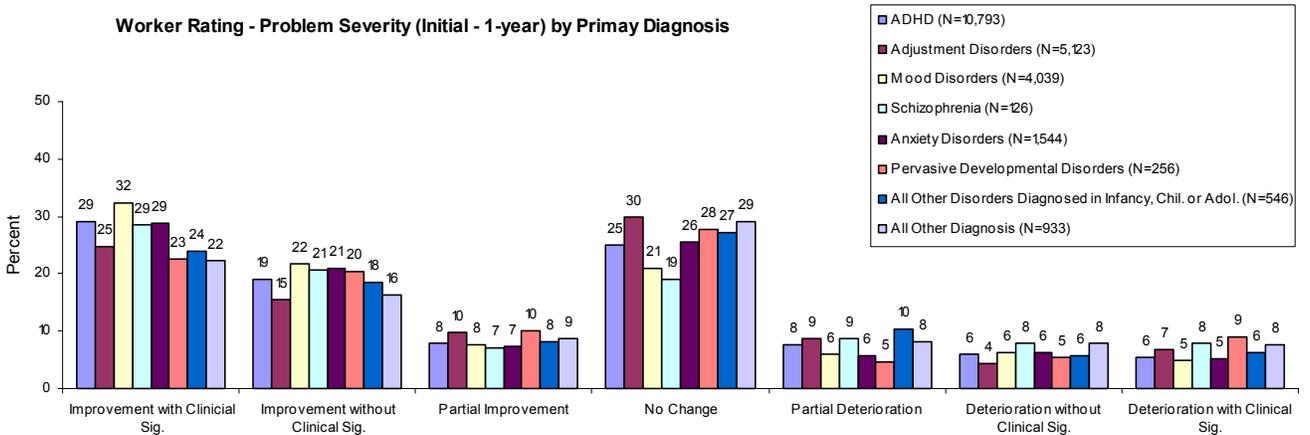
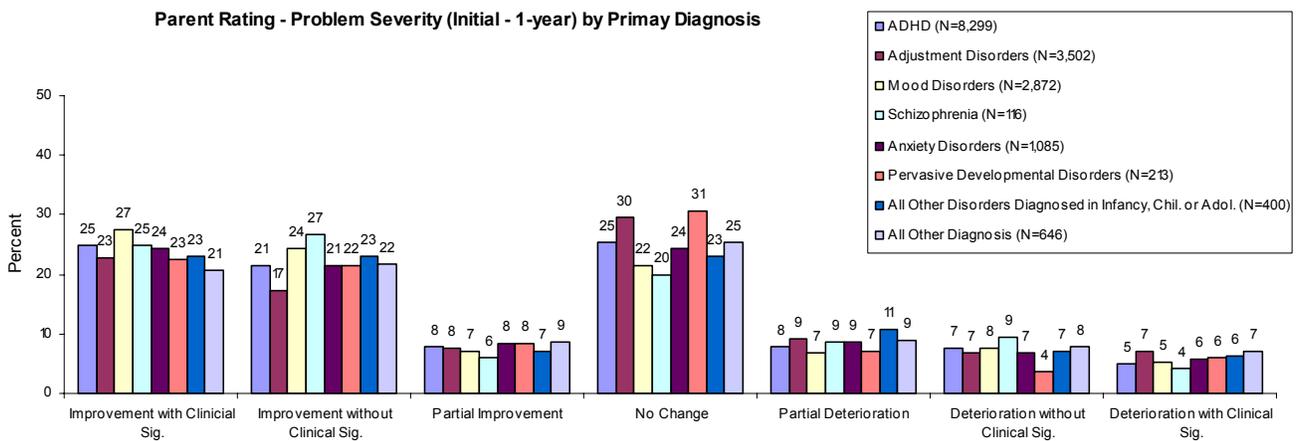
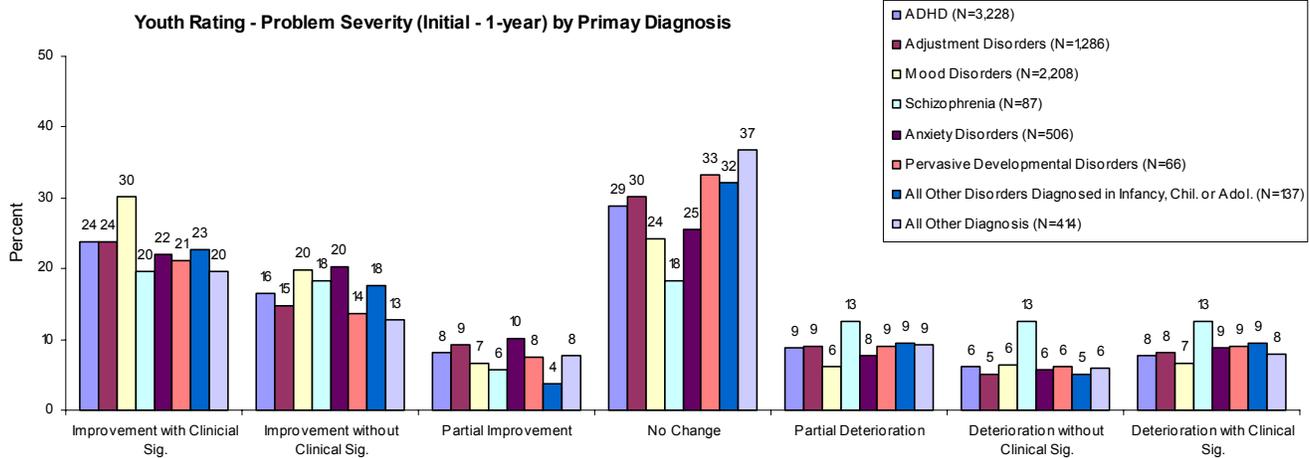
Change in Problem Severity by Primary Diagnosis, Initial Assessment to 3-month Assessment



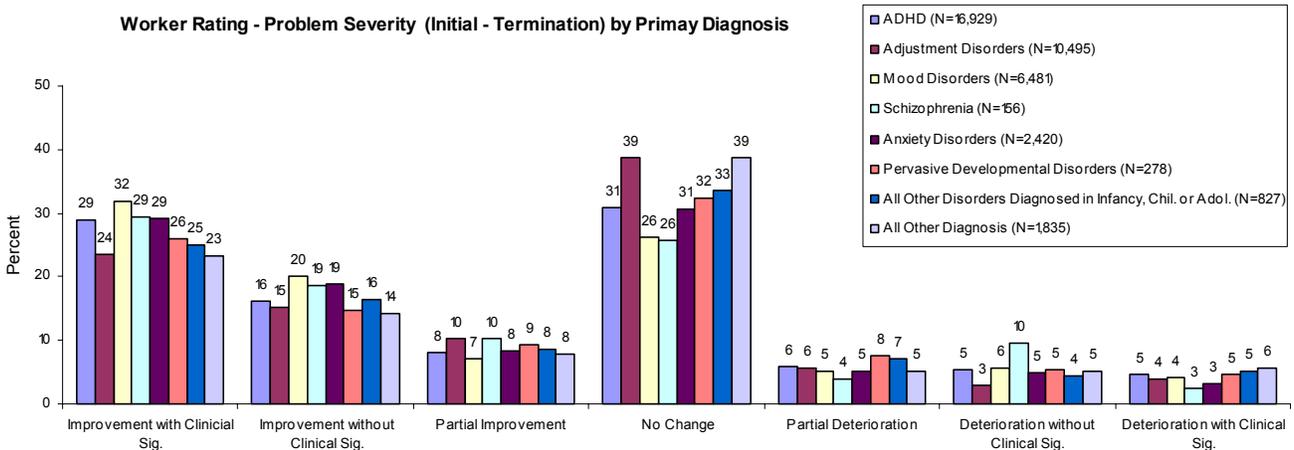
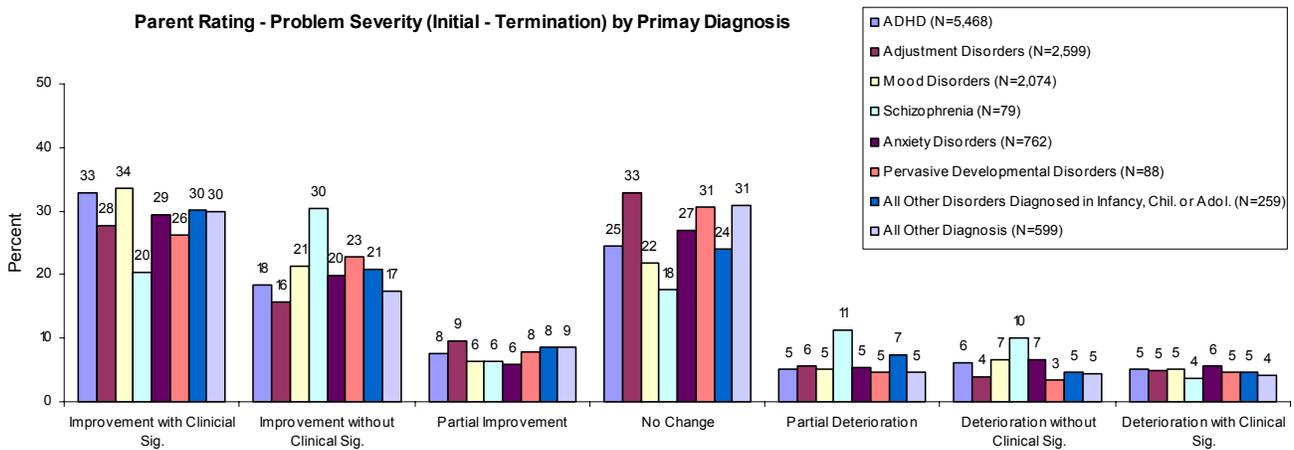
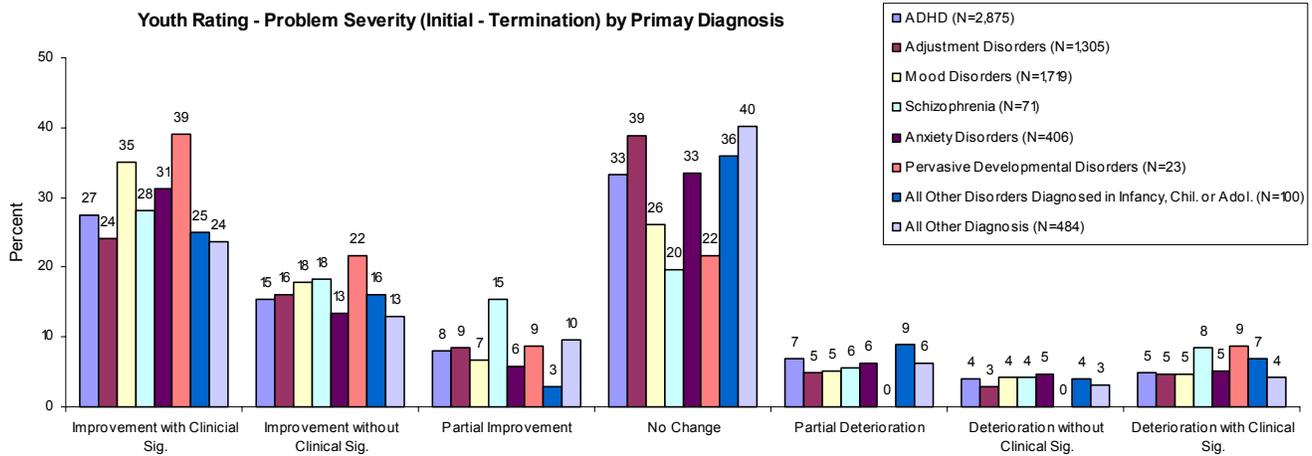
Change in Problem Severity by Primary Diagnosis, Initial Assessment to 6-month Assessment



Change in Problem Severity by Primary Diagnosis, Initial Assessment to 1-year Assessment



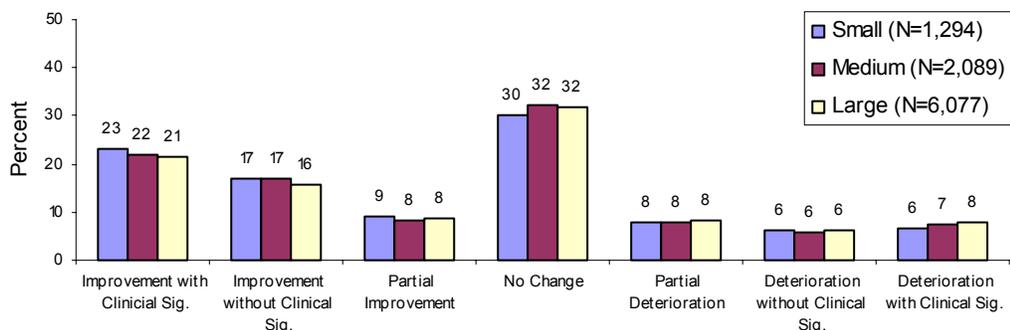
Change in Problem Severity by Primary Diagnosis, Initial Assessment to Termination



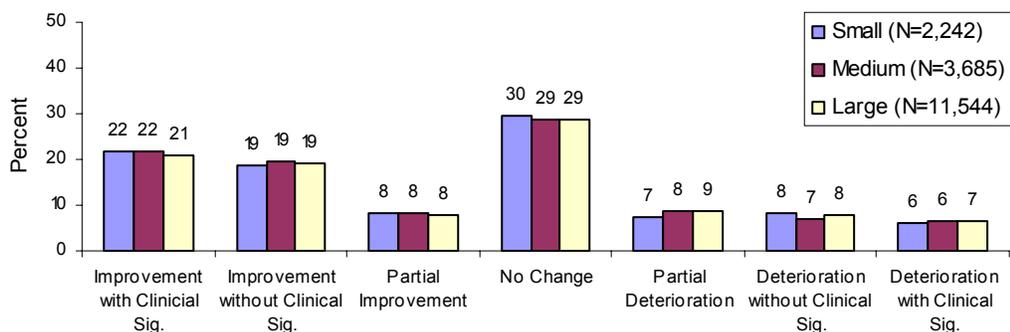
Changes by County Size—Problem Severity

Change in Problem Severity by County Size⁴, Initial Assessment to 3-month Assessment

Youth Rating - Problem Severity (Initial - 3-month) by County Size

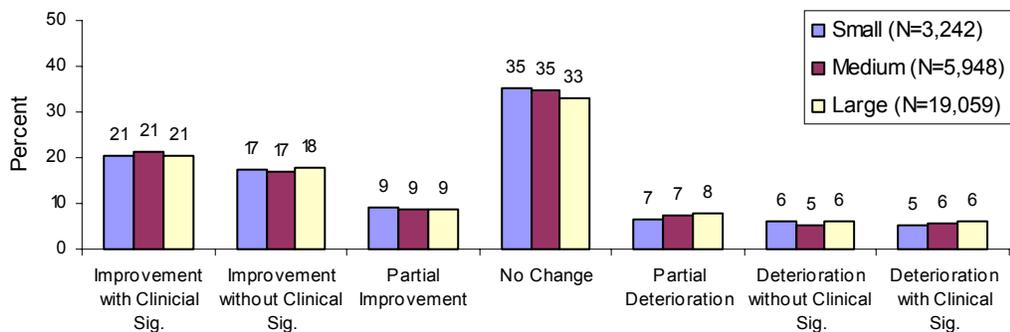


Parent Rating - Problem Severity (Initial - 3-month) by County Size



Note: County size is a significant factor in multivariate analysis with Problem Severity as dependent variable.

Worker Rating - Problem Severity (Initial - 3-month) by County Size

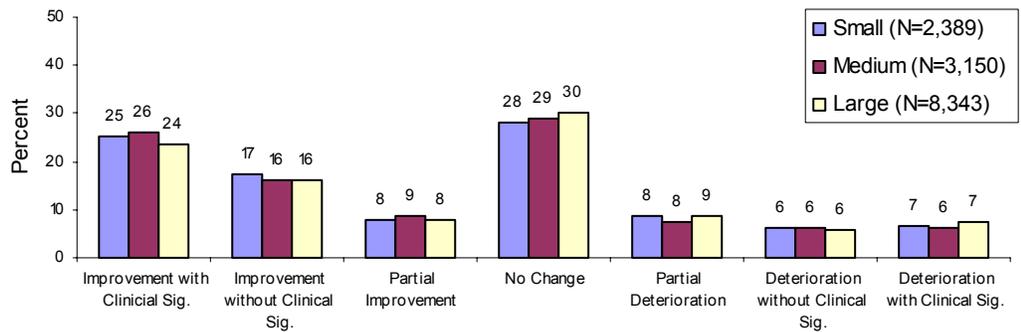


Note: County size is a significant factor in multivariate analysis with Problem Severity as dependent variable.

⁴ Large counties here refer to population greater than 300,000. Medium size counties will have a population between 100,000 and 300,000. Counties with population less than 100,000 are classified as small.

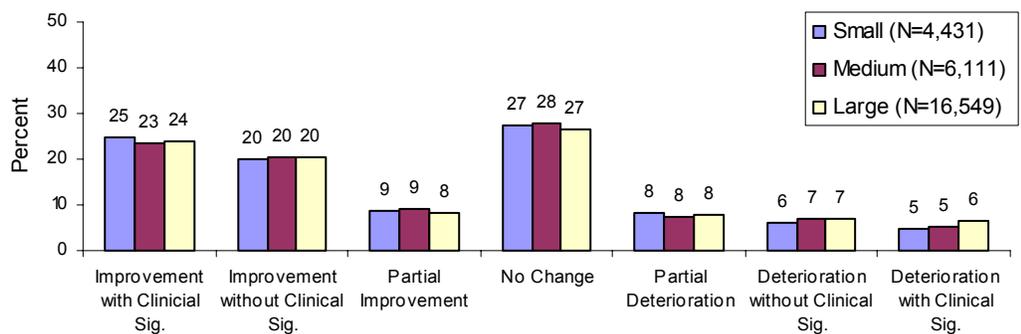
Change in Problem Severity by County Size, Initial Assessment to 6-month Assessment

Youth Rating - Problem Severity (Initial - 6-month) by County Size

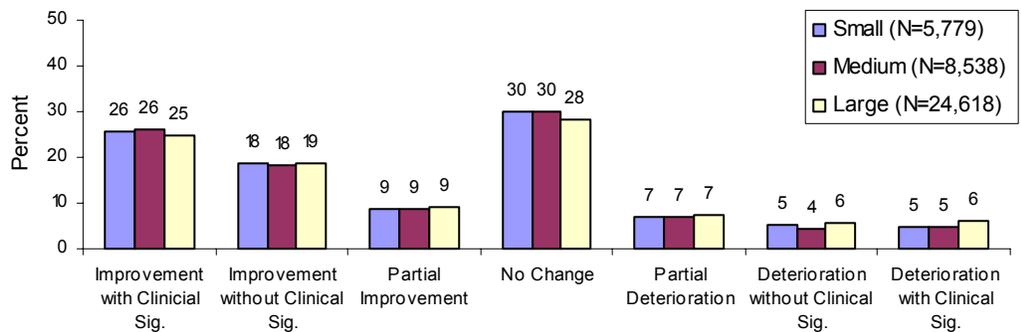


Note: County size is a significant factor in multivariate analysis with Problem Severity as dependent variable.

Parent Rating - Problem Severity (Initial - 6-month) by County Size



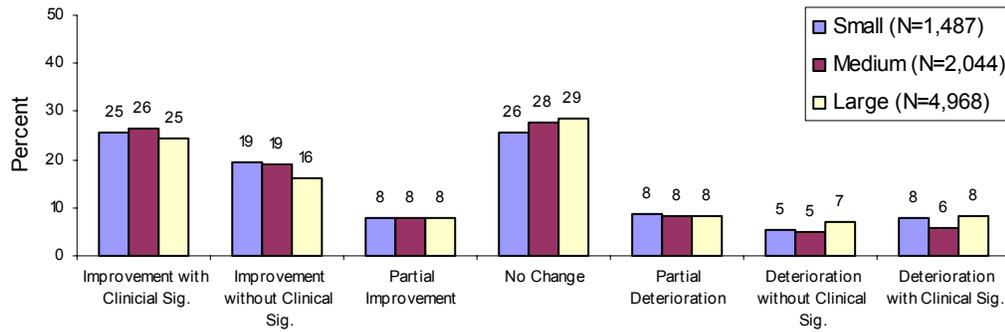
Worker Rating - Problem Severity (Initial - 6-month) by County Size



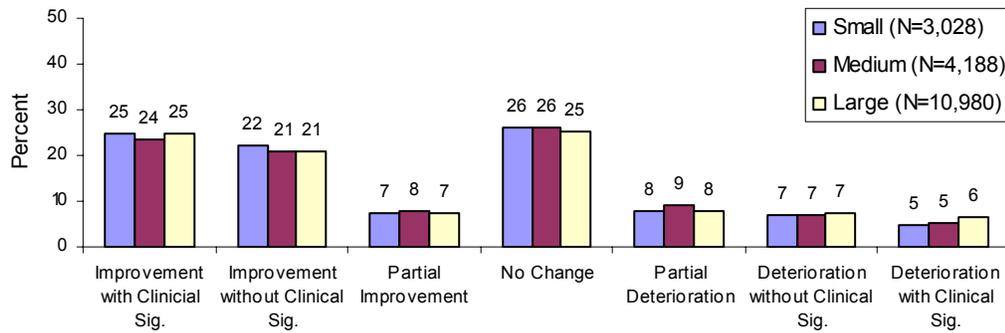
Note: County size is a significant factor in multivariate analysis with Problem Severity as dependent variable.

Change in Problem Severity by County Size, Initial Assessment to 1-year Assessment

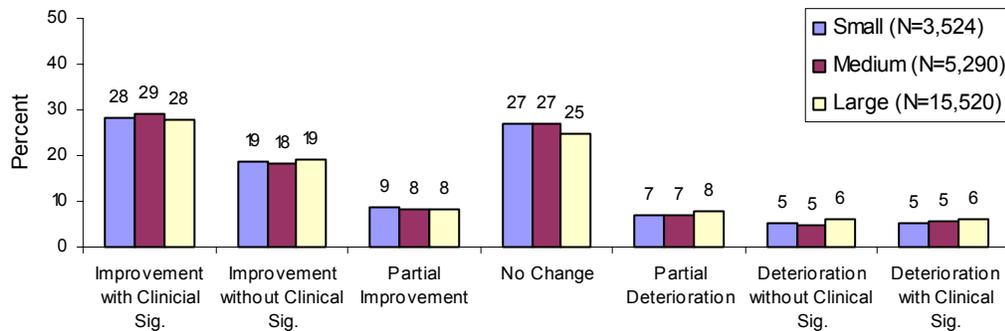
Youth Rating - Problem Severity (Initial - 1-year) by County Size



Parent Rating - Problem Severity (Initial - 1-year) by County Size

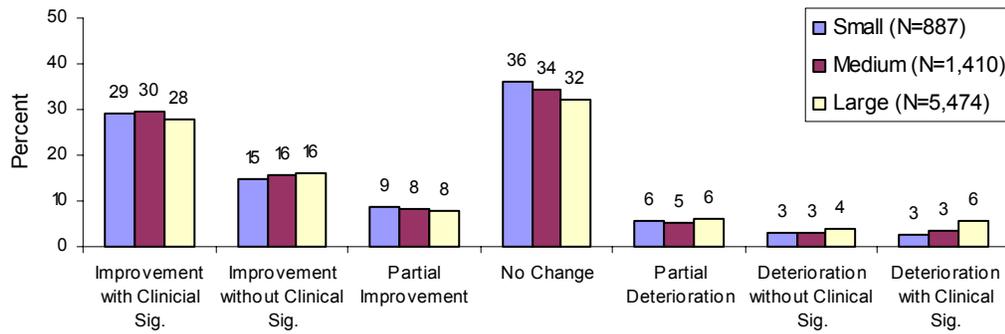


Worker Rating - Problem Severity (Initial - 1-year) by County Size

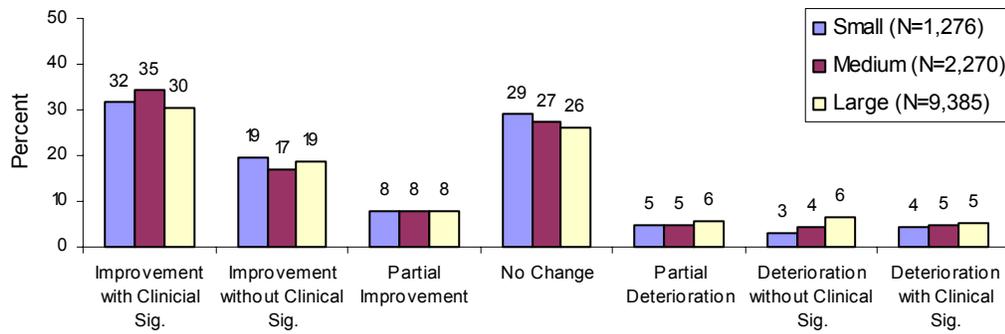


Change in Problem Severity by County Size, Initial Assessment to Termination

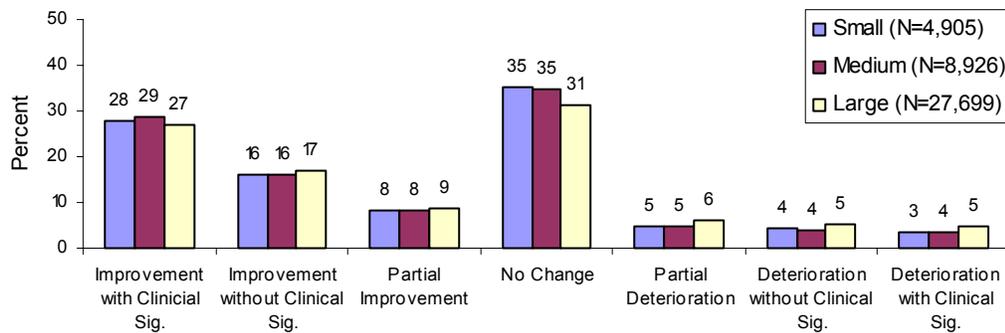
Youth Rating - Problem Severity (Initial - Termination) by County Size



Parent Rating - Problem Severity (Initial - Termination) by County Size

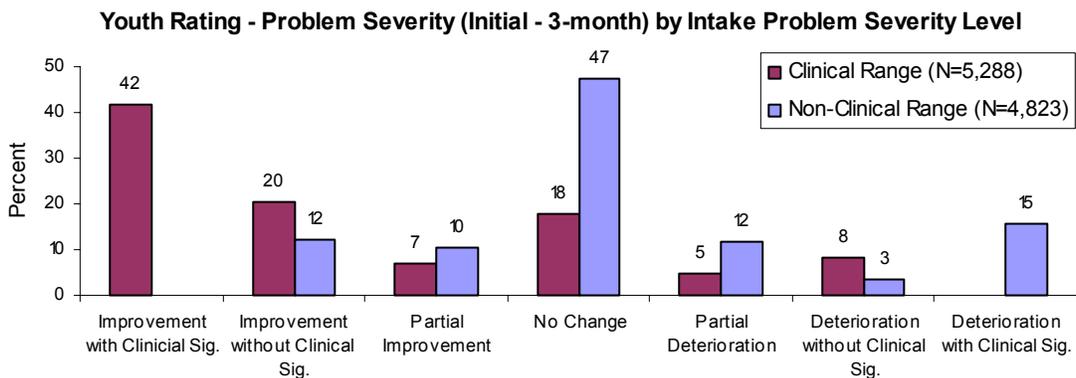


Woker Rating - Problem Severity (Initial - Termination) by County Size

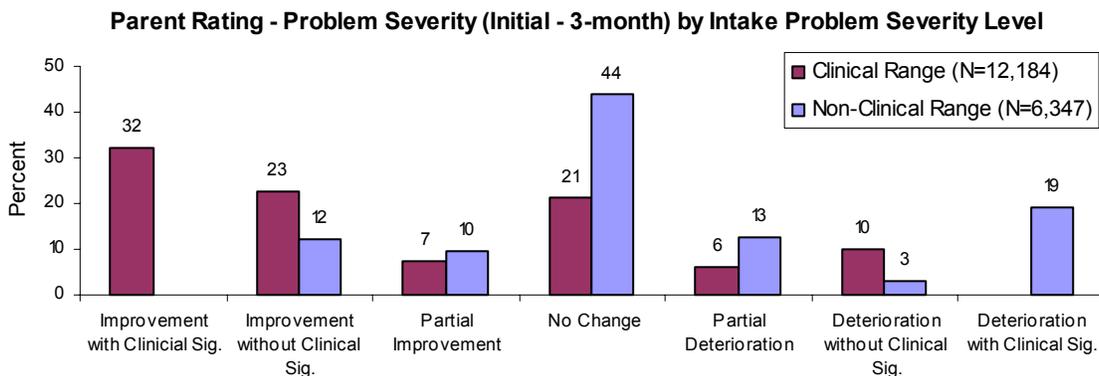


Changes by Intake Problem Severity—Problem Severity

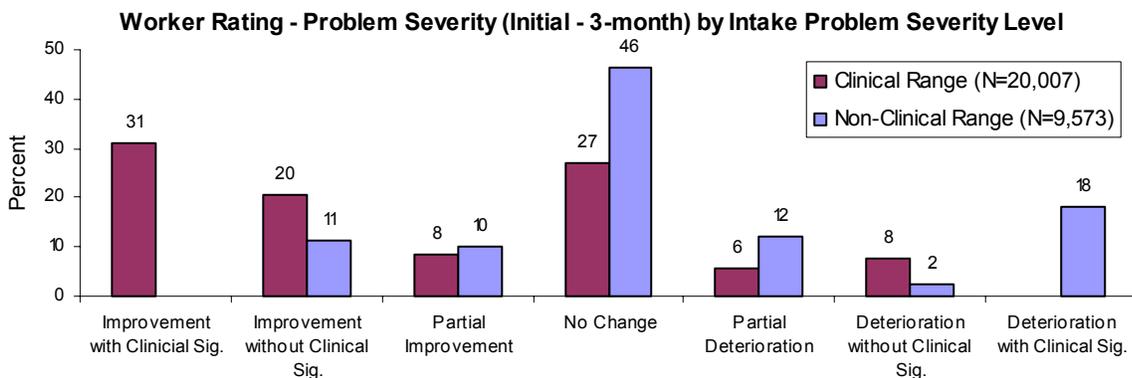
Change in Problem Severity by Intake Problem Severity Level, Initial Assessment to 3-month Assessment



Note: Intake Problem Severity is a significant factor in multivariate analysis with Problem Severity as dependent variable.

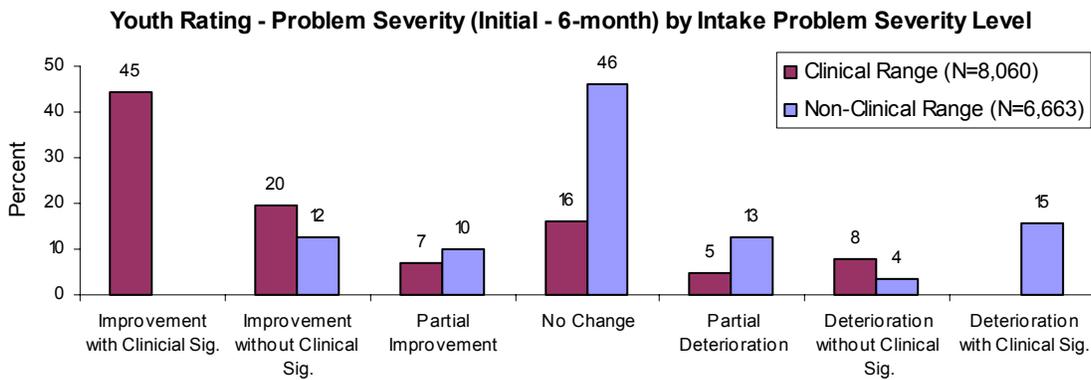


Note: Intake Problem Severity is a significant factor in multivariate analysis with Problem Severity as dependent variable.

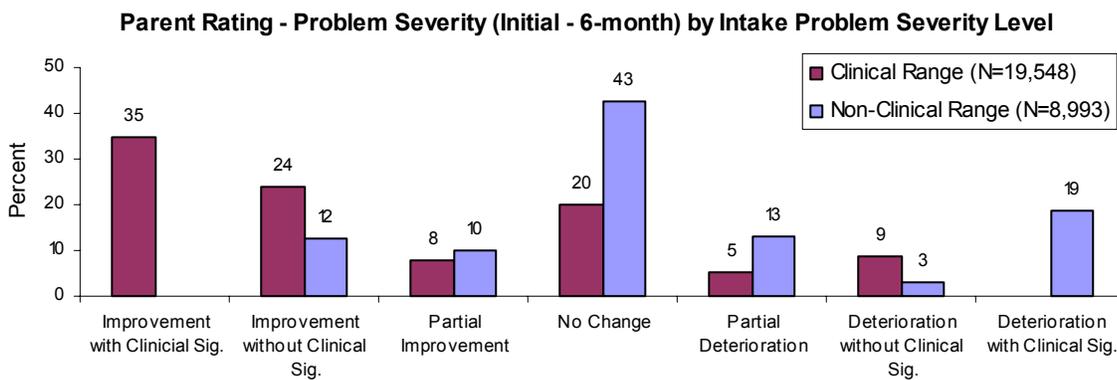


Note: Intake Problem Severity is a significant factor in multivariate analysis with Problem Severity as dependent variable.

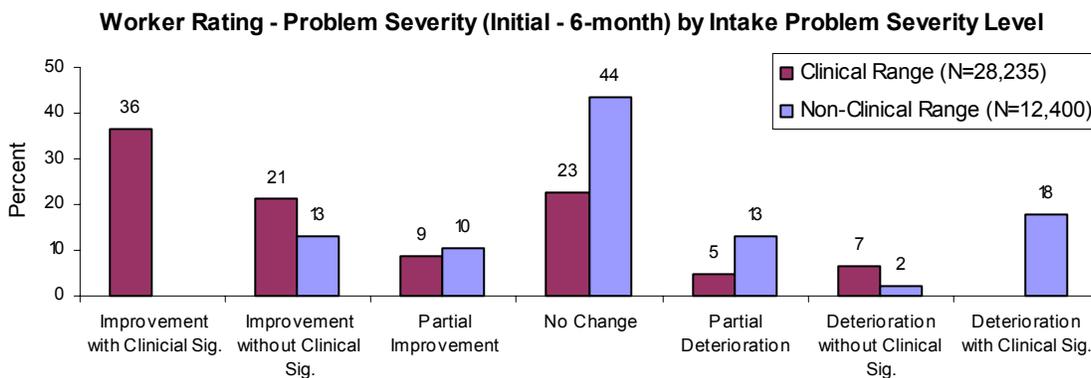
Change in Problem Severity by Intake Problem Severity Level, Initial Assessment to 6-month Assessment



Note: Intake Problem Severity is a significant factor in multivariate analysis with Problem Severity as dependent variable.

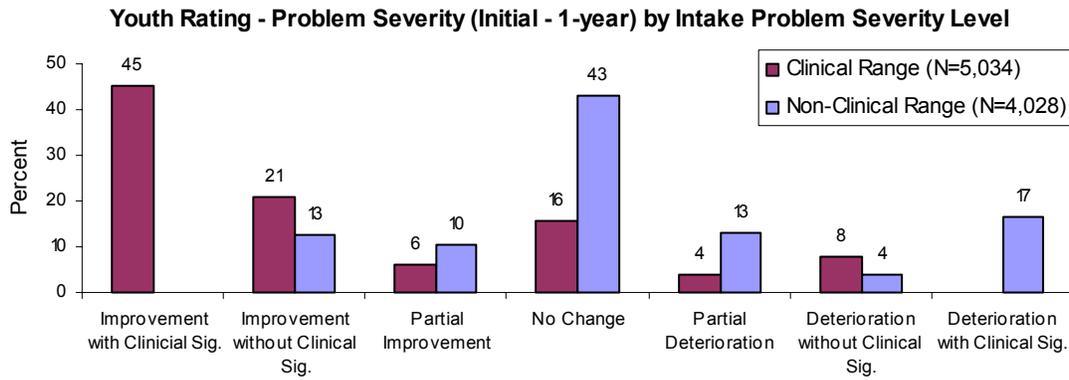


Note: Intake Problem Severity is a significant factor in multivariate analysis with Problem Severity as dependent variable.

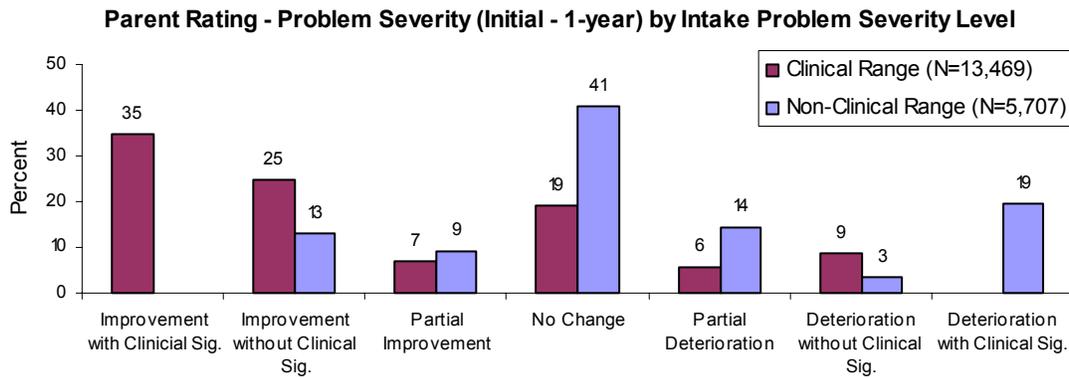


Note: Intake Problem Severity is a significant factor in multivariate analysis with Problem Severity as dependent variable.

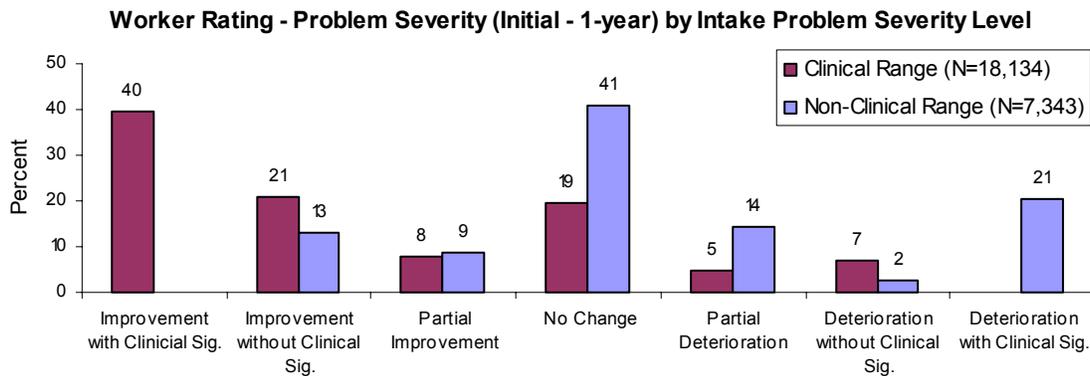
Change in Problem Severity by Intake Problem Severity Level, Initial Assessment to 1-year Assessment



Note: Intake Problem Severity is a significant factor in multivariate analysis with Problem Severity as dependent variable.

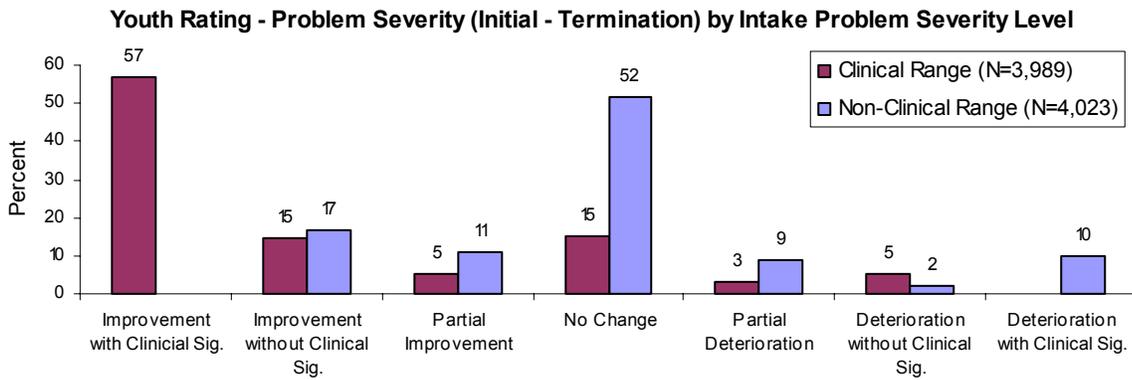


Note: Intake Problem Severity is a significant factor in multivariate analysis with Problem Severity as dependent variable.

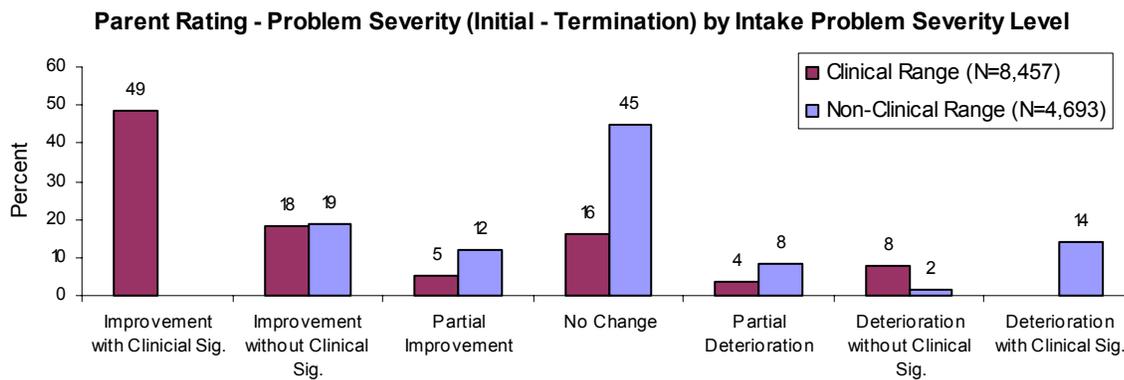


Note: Intake Problem Severity is a significant factor in multivariate analysis with Problem Severity as dependent variable.

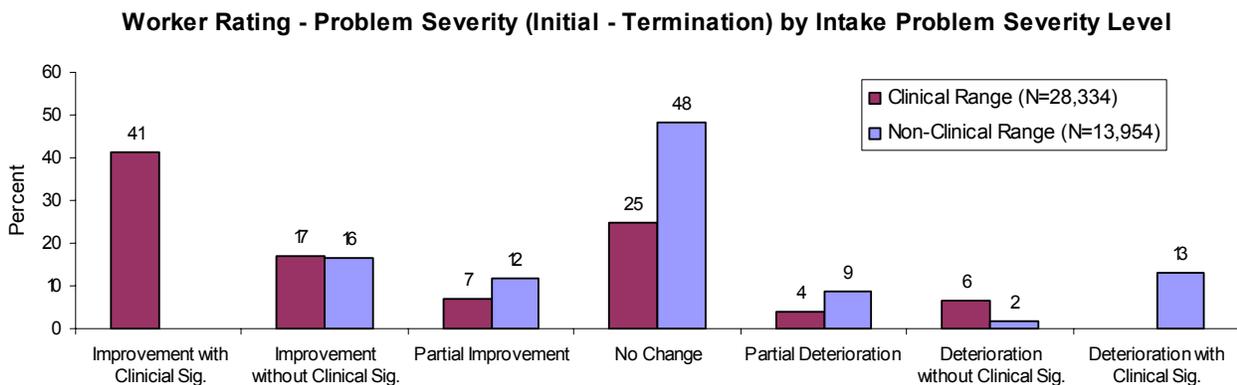
Change in Problem Severity by Intake Problem Severity Level, Initial Assessment to Termination



Note: Intake Problem Severity is a significant factor in multivariate analysis with Problem Severity as dependent variable.



Note: Intake Problem Severity is a significant factor in multivariate analysis with Problem Severity as dependent variable.



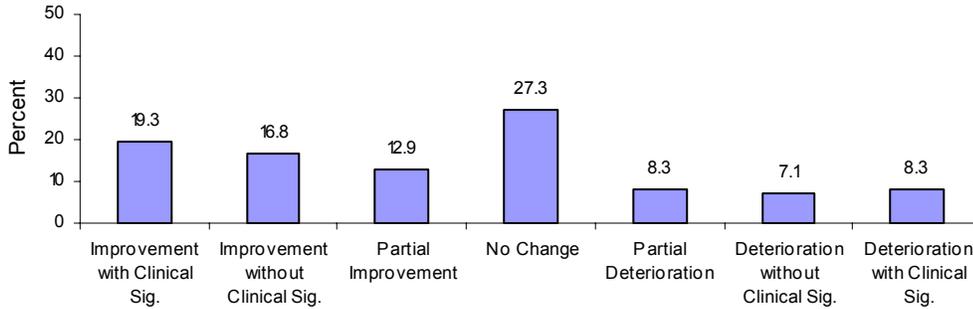
Note: Intake Problem Severity is a significant factor in multivariate analysis with Problem Severity as dependent variable.

Change in Ohio Scales - Functioning

Overall Changes—Functioning

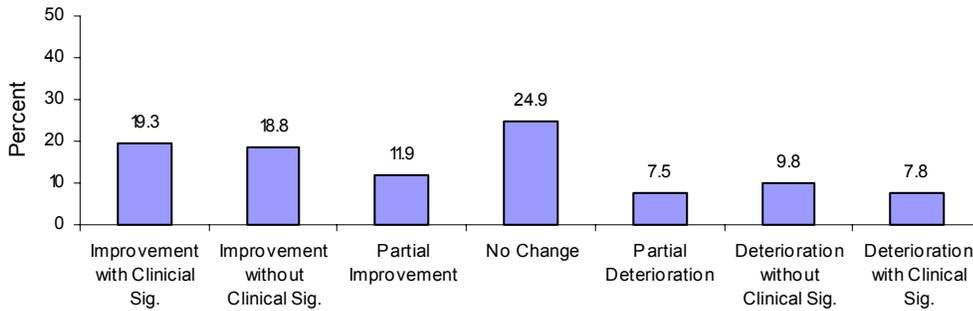
Overall change in Functioning, Initial Assessment to 3-month Assessment

Youth Rating - Functioning (Initial - 3-month) (N=10,149)



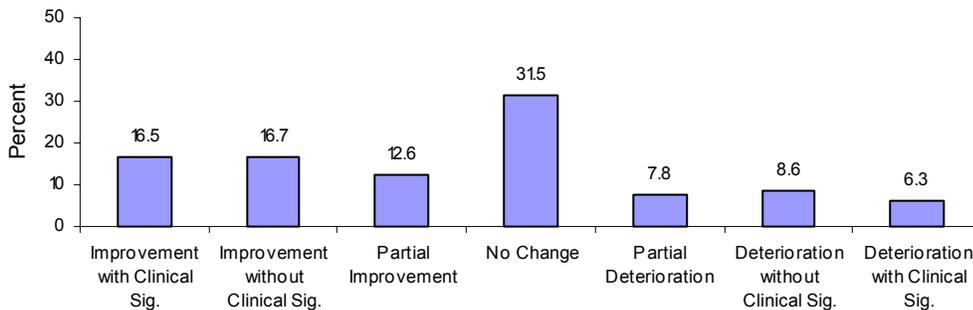
Note: Average number of days between assessments = 94 (S.D.=23)

Parent Rating - Functioning (Initial - 3-month) (N=18,458)



Note: Average number of days between assessments = 94 (S.D.=23)

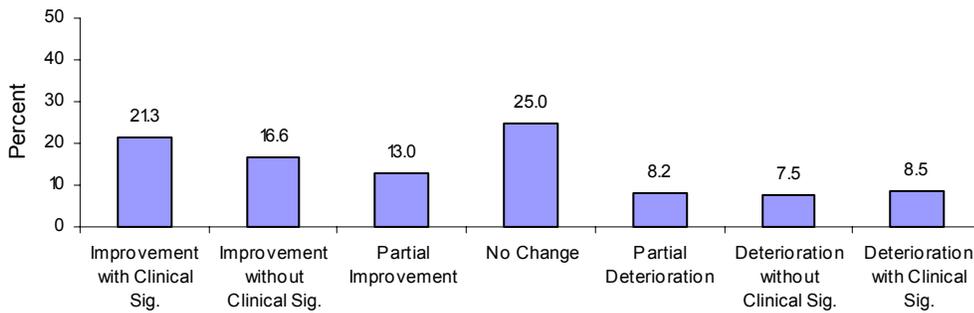
Worker Rating - Functioning (Initial - 3-month) (N=29,398)



Note: Average number of days between assessments = 93 (S.D.=23)

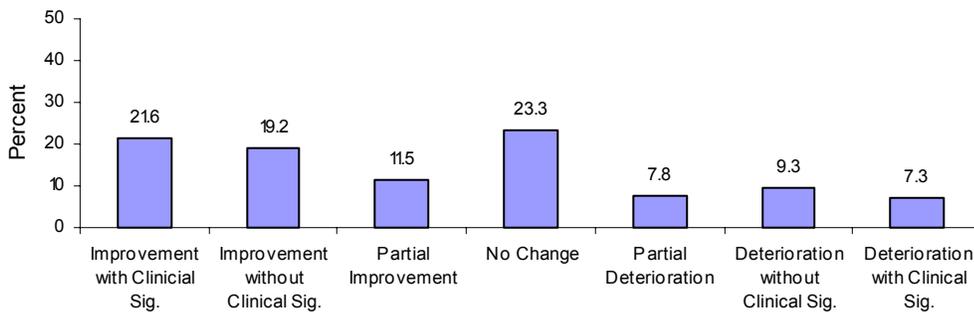
Overall change in Functioning, Initial Assessment to 6-month Assessment

Youth Rating - Functioning (Initial - 6-month) (N=14,756)



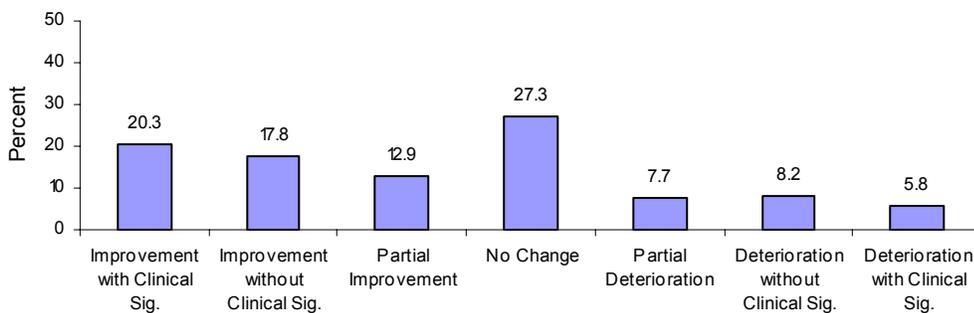
Note: Average number of days between assessments = 194 (S.D.=33)

Parent Rating - Functioning (Initial - 6-month) (N=28,265)



Note: Average number of days between assessments = 194 (S.D.=32)

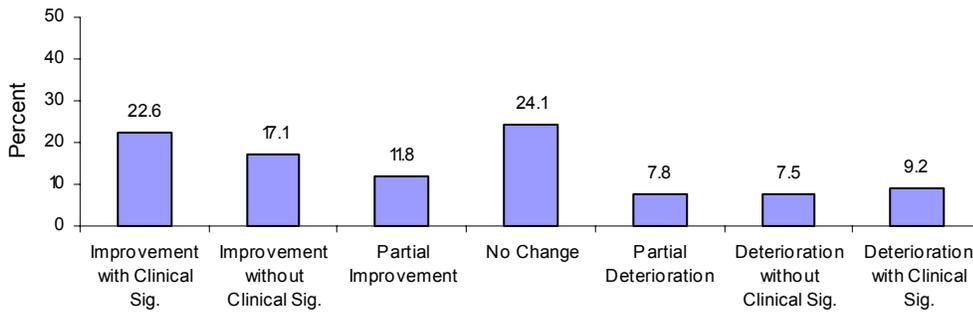
Worker Rating - Functioning (Initial - 6-month) (N=40,369)



Note: Average number of days between assessments = 193 (S.D.=33)

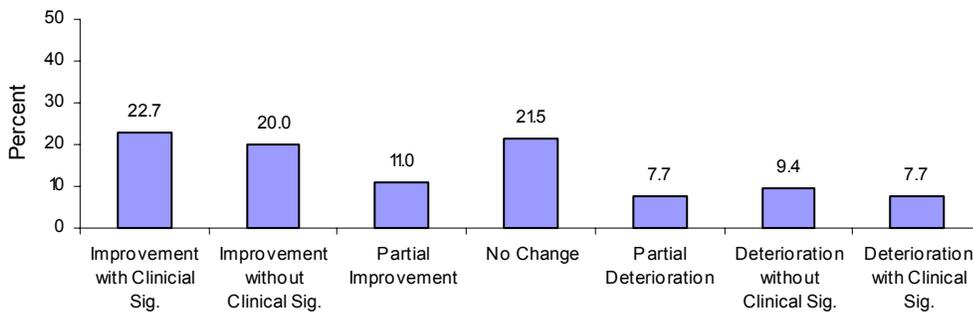
Overall change in Functioning, Initial Assessment to 1-year Assessment

Youth Rating - Functioning (Initial - 1-year) (N=9,048)



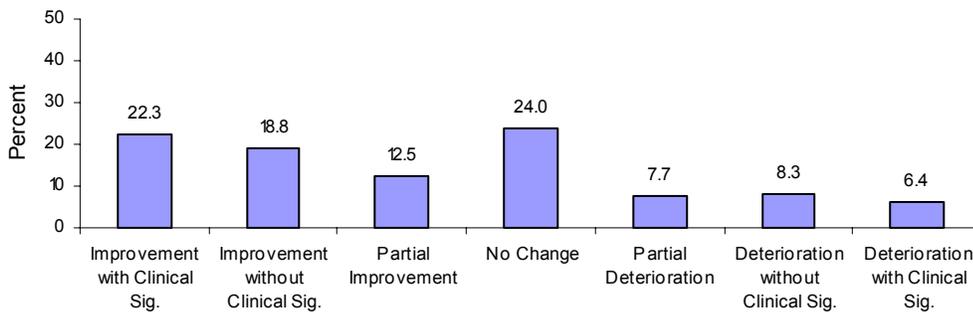
Note: Average number of days between assessments = 355 (S.D.=46)

Parent Rating - Functioning (Initial - 1-year) (N=18,902)



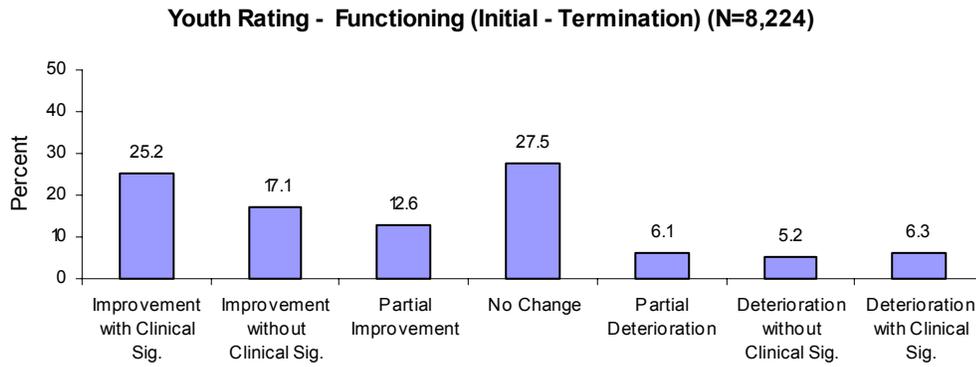
Note: Average number of days between assessments = 356 (S.D.=46)

Worker Rating - Functioning (Initial - 1-year) (N=25,250)

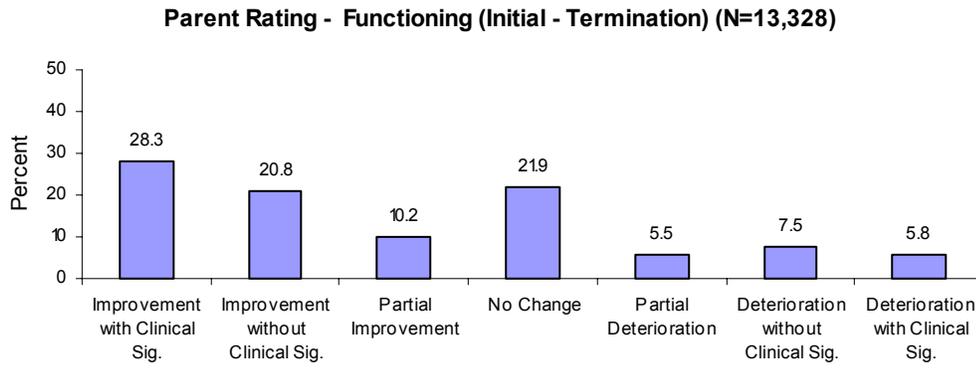


Note: Average number of days between assessments = 352 (S.D.=46)

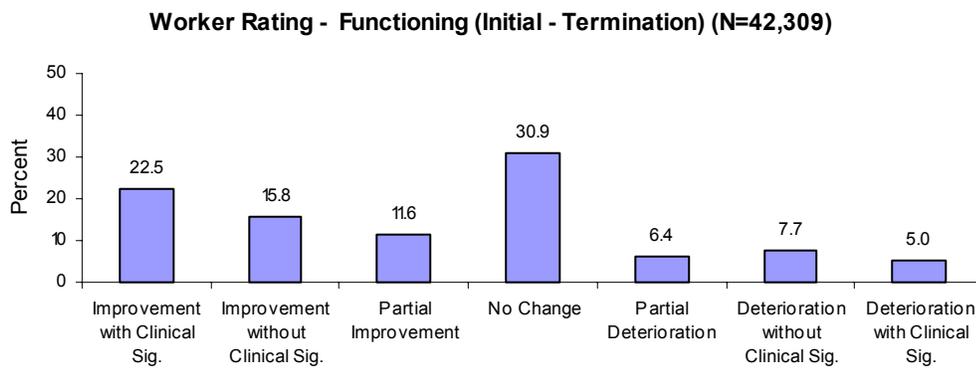
Overall change in Functioning, Initial Assessment to Termination



Note: Average number of days between assessments = 303 (S.D.=390)



Note: Average number of days between assessments = 329 (S.D.=378)

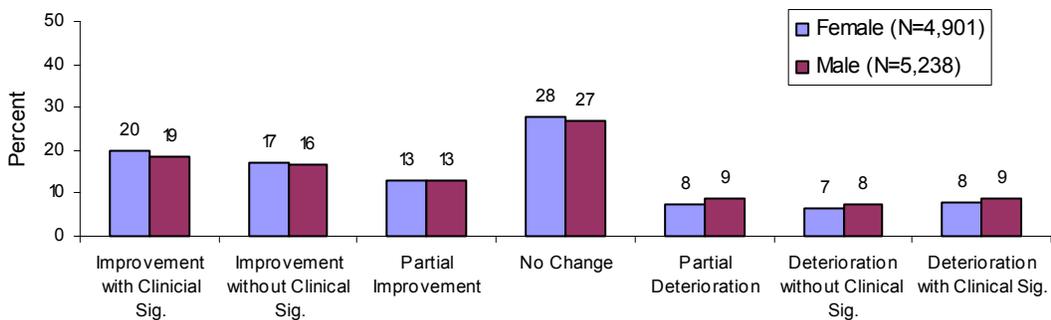


Note: Average number of days between assessments = 299 (S.D.=343)

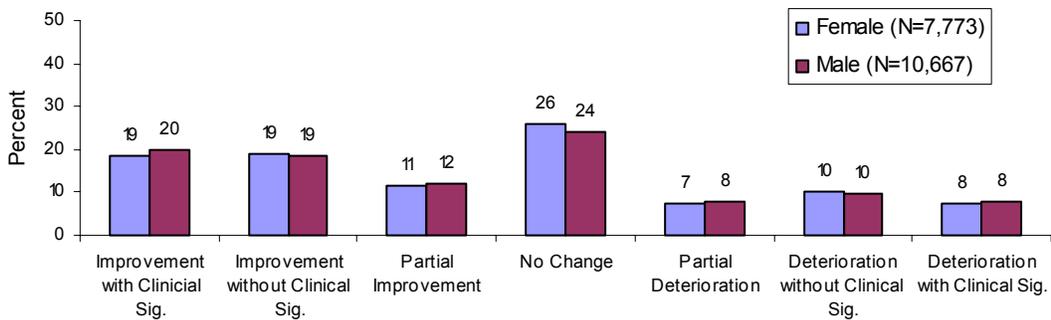
Changes by Gender—Functioning

Change in Functioning by Gender, Initial Assessment to 3-month Assessment

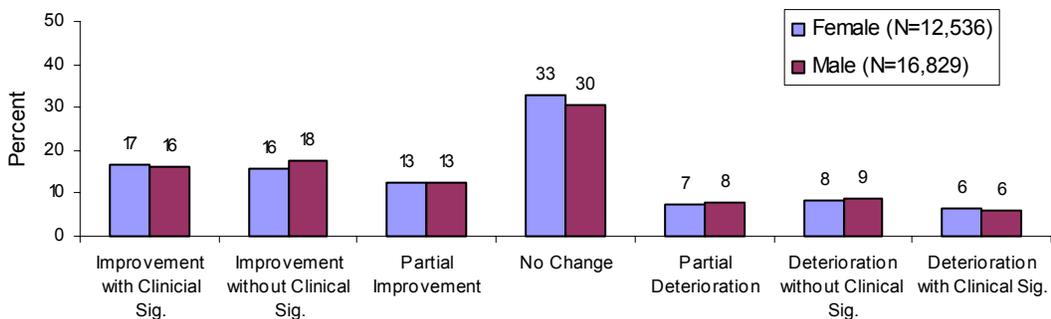
Youth Rating - Functioning (Initial - 3-month) by Gender



Parent Rating - Functioning (Initial - 3-month) by Gender

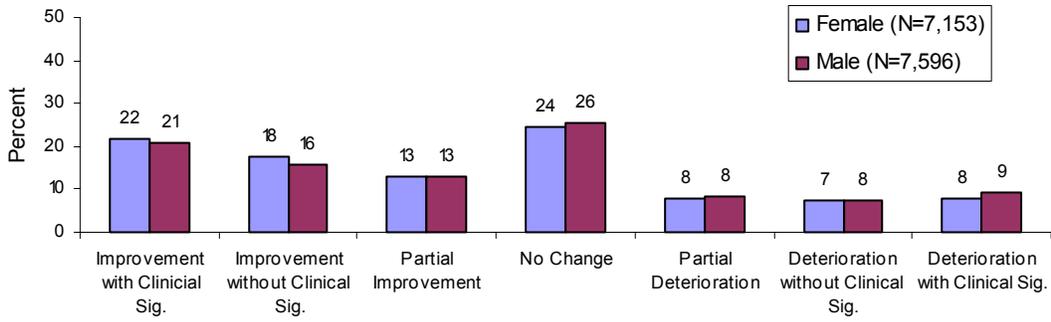


Worker Rating - Functioning (Initial - 3-month) by Gender

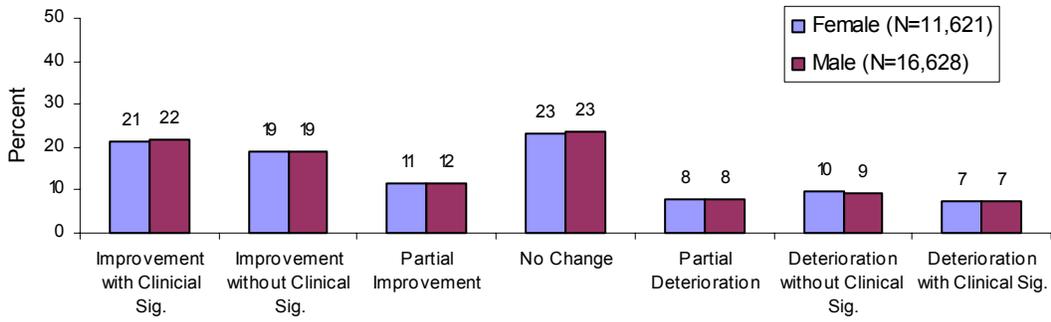


Change in Functioning by Gender, Initial Assessment to 6-month Assessment

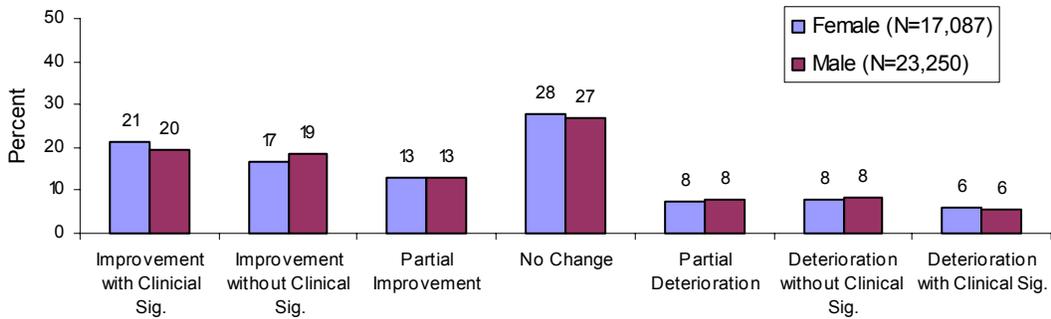
Youth Rating - Functioning (Initial - 6-month) by Gender



Parent Rating - Functioning (Initial - 6-month) by Gender

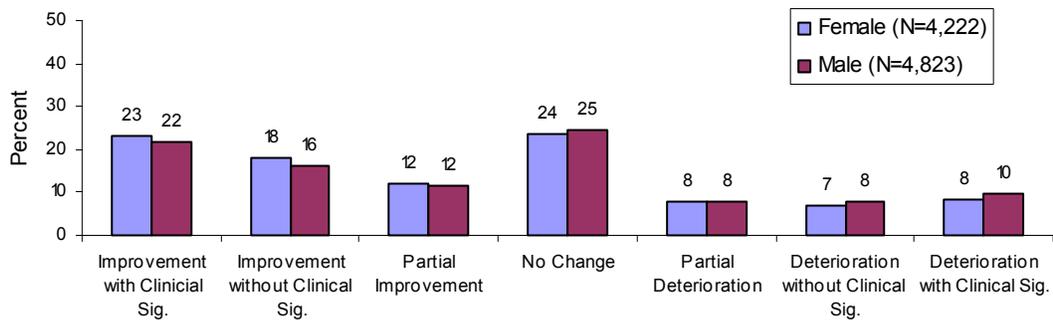


Worker Rating - Functioning (Initial - 6-month) by Gender

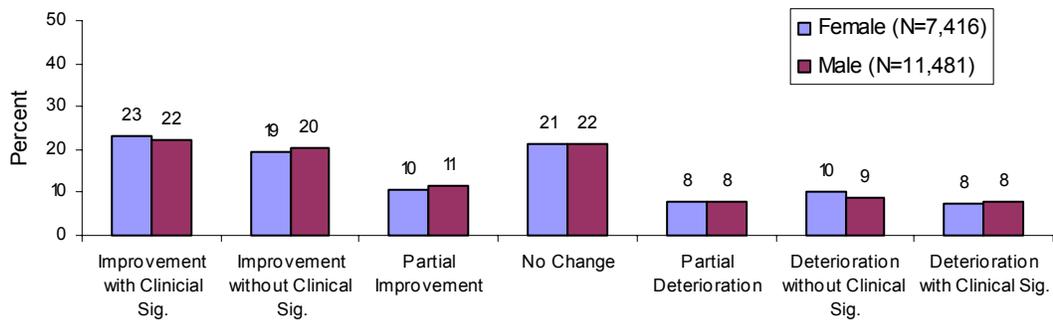


Change in Functioning by Gender, Initial Assessment to 1-year Assessment

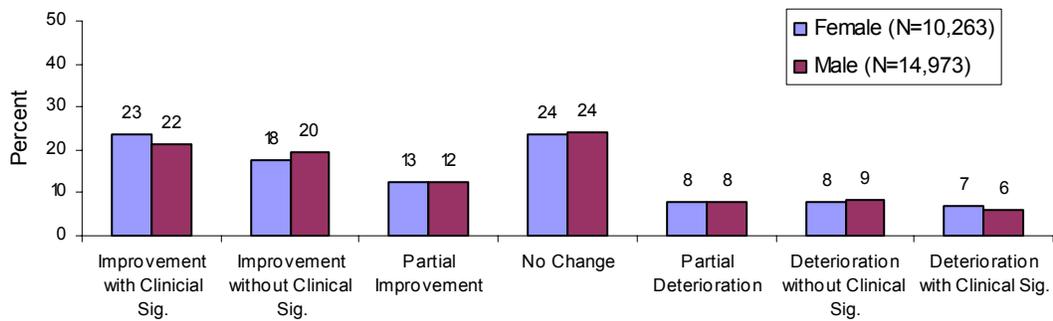
Youth Rating - Functioning (Initial - 1-year) by Gender



Parent Rating - Functioning (Initial - 1-year) by Gender

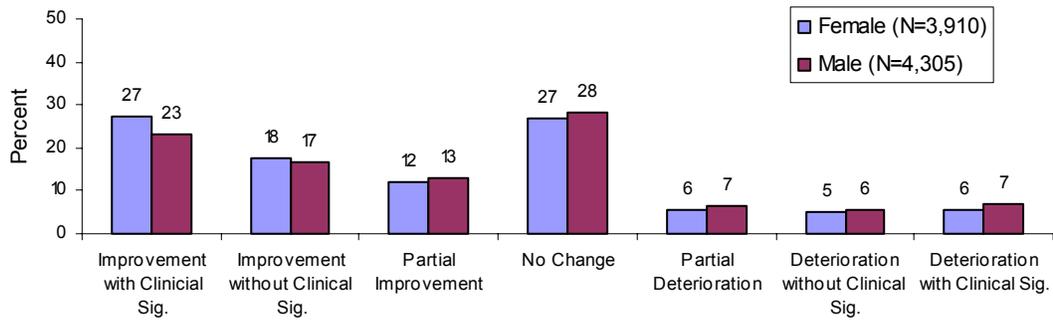


Worker Rating - Functioning (Initial - 1-year) by Gender

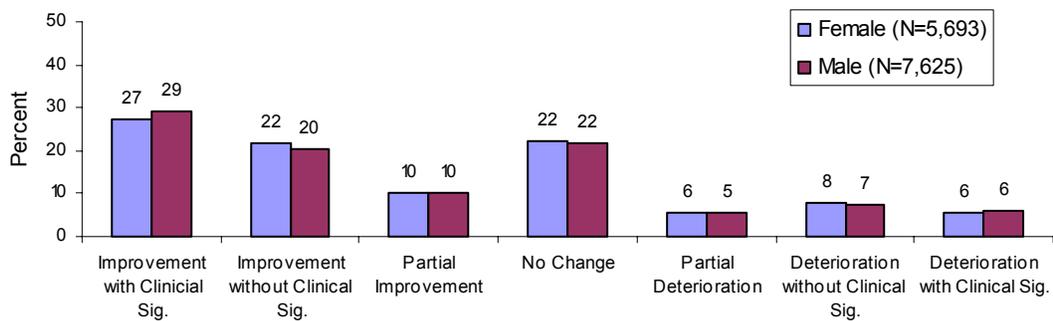


Change in Functioning by Gender, Initial Assessment to Termination

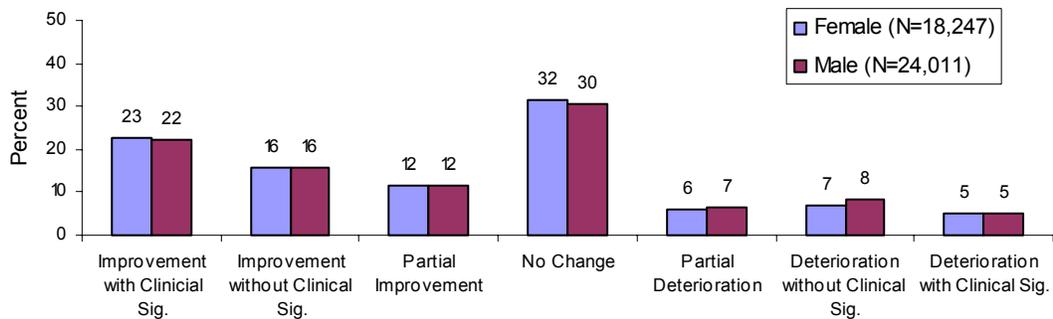
Youth Rating - Functioning (Initial - Termination) by Gender



Parent Rating - Functioning (Initial - Termination) by Gender



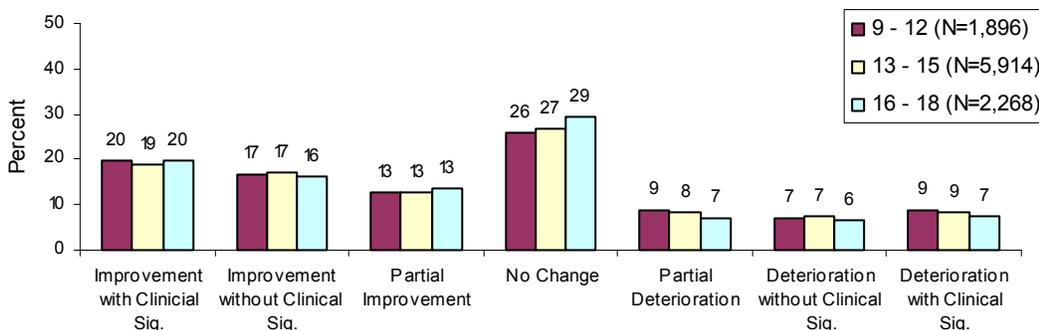
Worker Rating - Functioning (Initial - Termination) by Gender



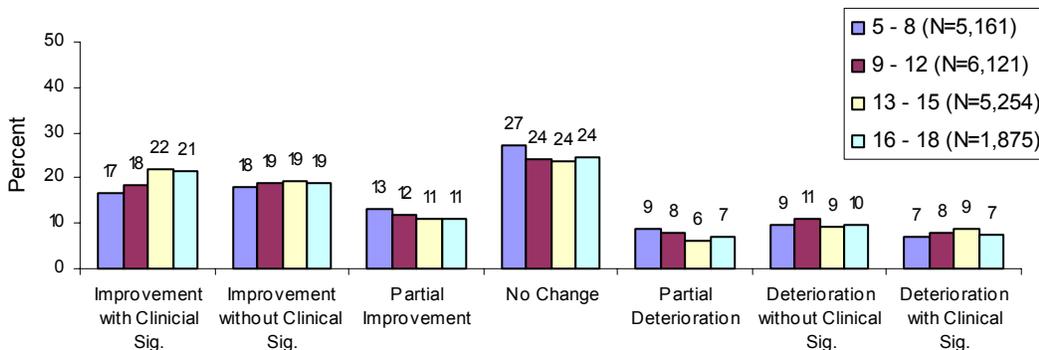
Changes by Age—Functioning

Change in Functioning by Age, Initial Assessment to 3-month Assessment

Youth Rating - Functioning (Initial - 3-month) by Age

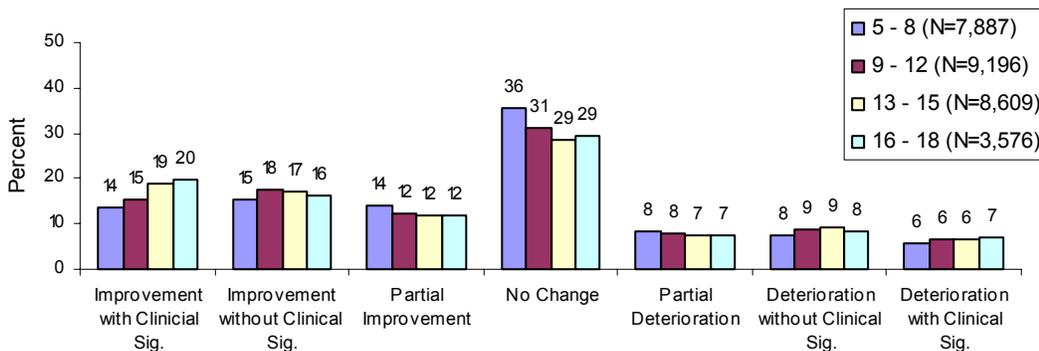


Parent Rating - Functioning (Initial - 3-month) by Age



Note: Age is a significant covariate in multivariate analysis with Functioning as dependent variable.

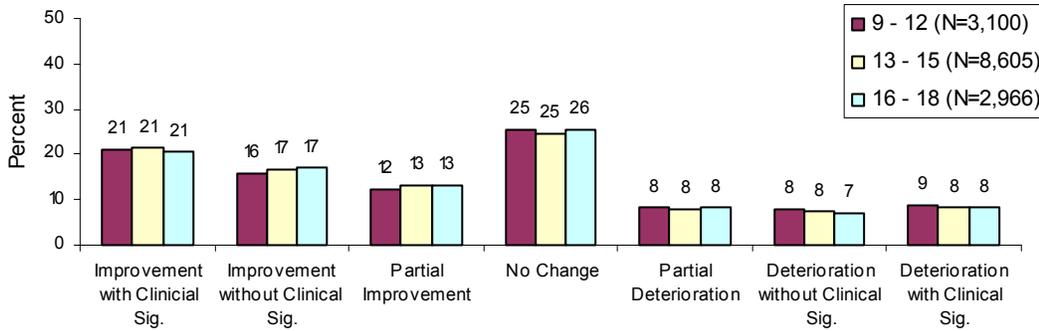
Worker Rating - Functioning (Initial - 3-month) by Age



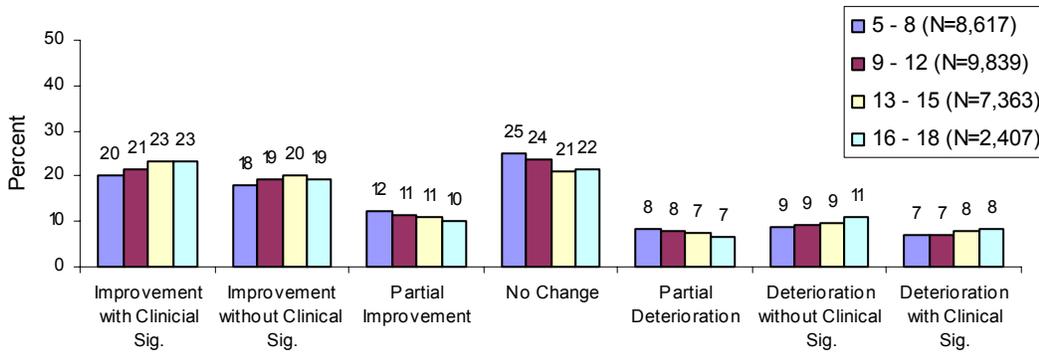
Note: Age is a significant covariate in multivariate analysis with Functioning as dependent variable.

Change in Functioning by Age, Initial Assessment to 6-month Assessment

Youth Rating - Functioning (Initial - 6-month) by Age

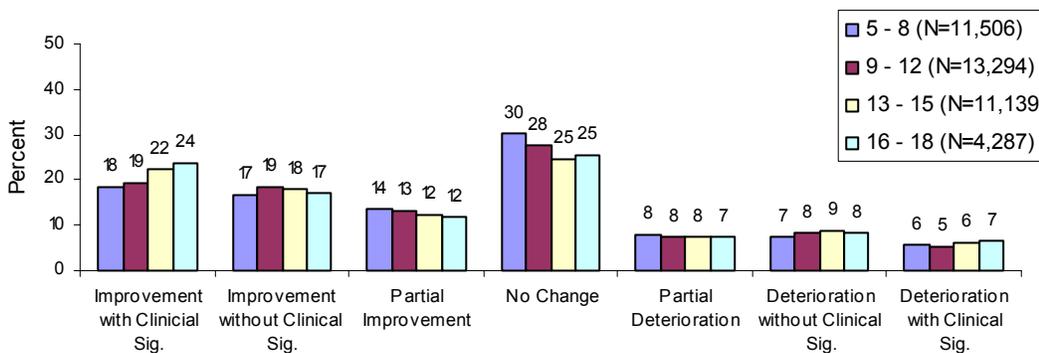


Parent Rating - Functioning (Initial - 6-month) by Age



Note: Age is a significant covariate in multivariate analysis with Functioning as dependent variable.

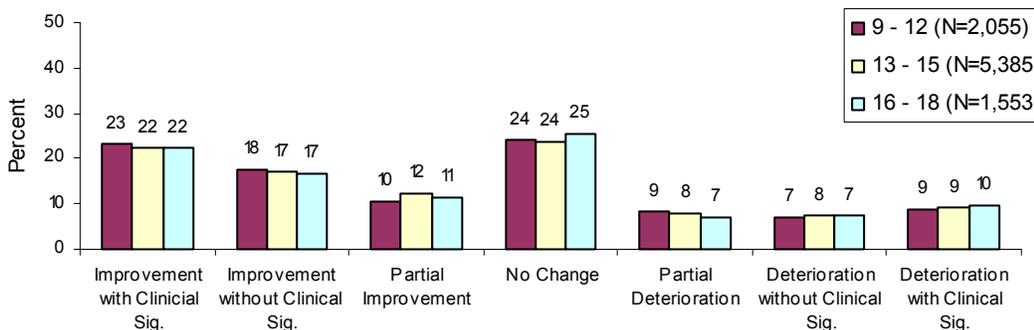
Worker Rating - Functioning (Initial - 6-month) by Age



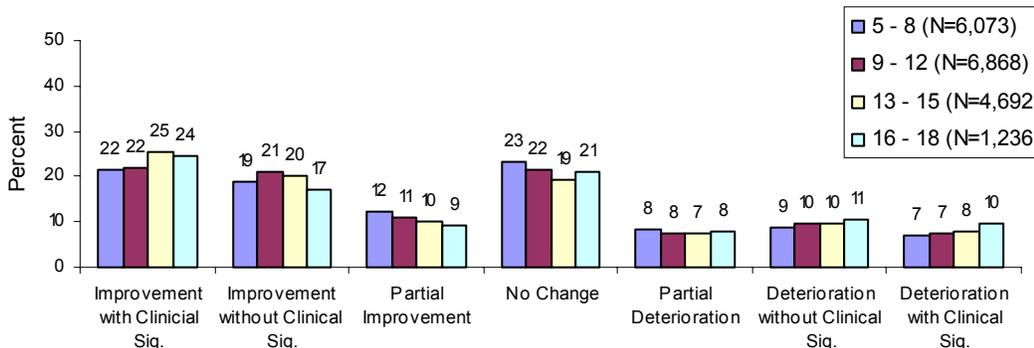
Note: Age is a significant covariate in multivariate analysis with Functioning as dependent variable.

Change in Functioning by Age, Initial Assessment to 1-year Assessment

Youth Rating - Functioning (Initial - 1-year) by Age

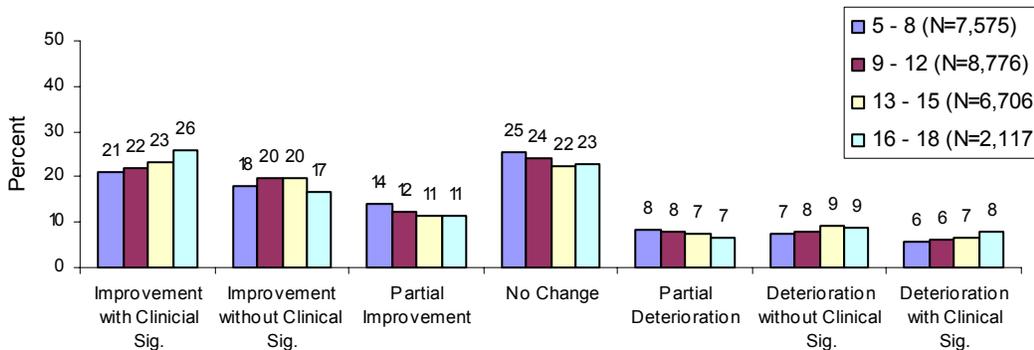


Parent Rating - Functioning (Initial - 1-year) by Age



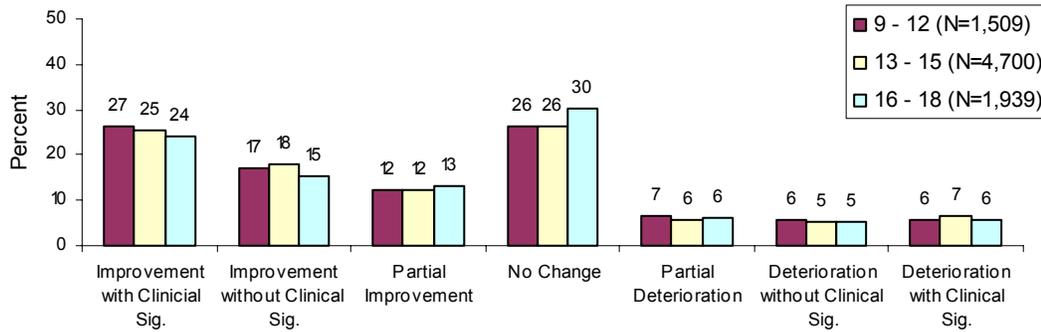
Note: Age is a significant covariate in multivariate analysis with Functioning as dependent variable.

Worker Rating - Functioning (Initial - 1-year) by Age

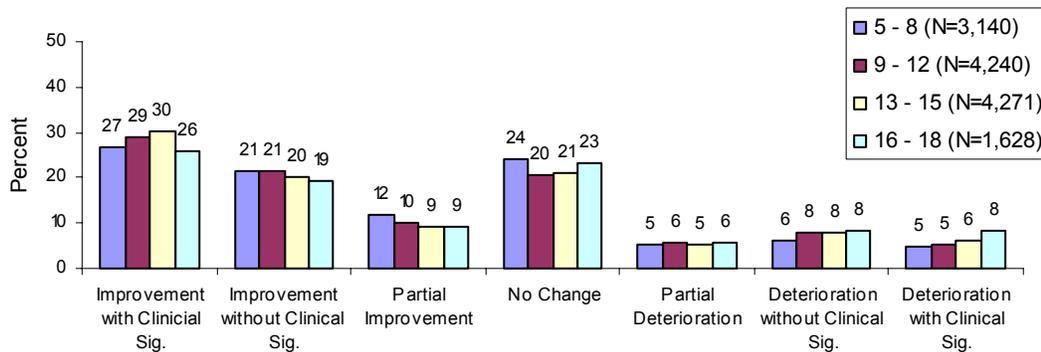


Change in Functioning by Age, Initial Assessment to Termination

Youth Rating - Functioning (Initial - Termination) by Age

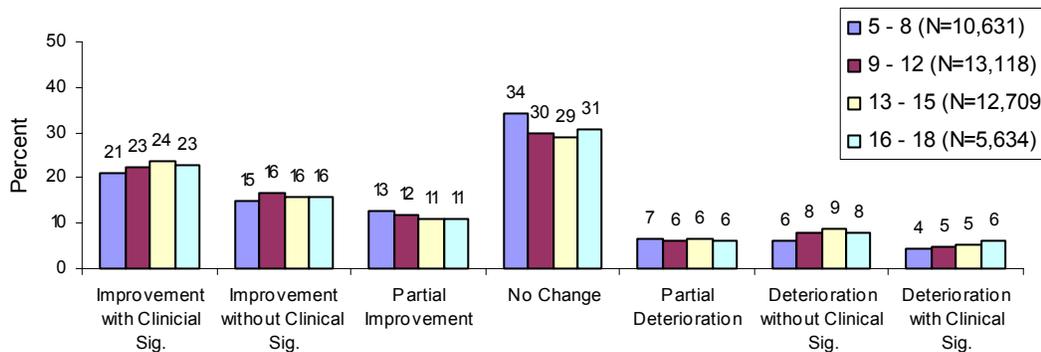


Parent Rating - Functioning (Initial - Termination) by Age



Note: Age is a significant covariate in multivariate analysis with Functioning as dependent variable.

Worker Rating - Functioning (Initial - Termination) by Age

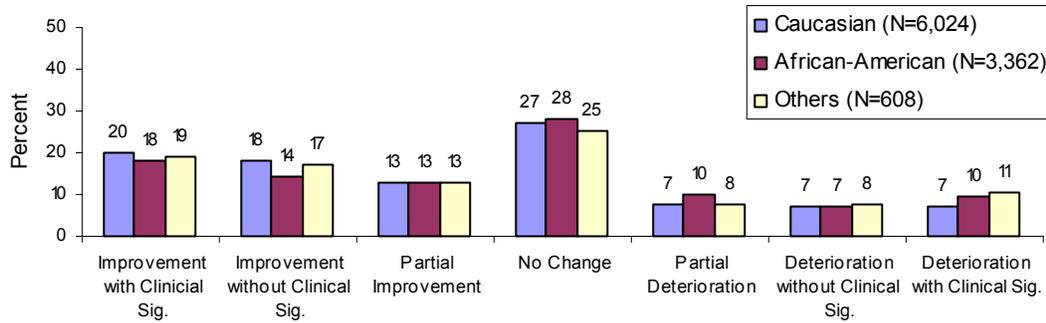


Note: Age is a significant covariate in multivariate analysis with Functioning as dependent variable.

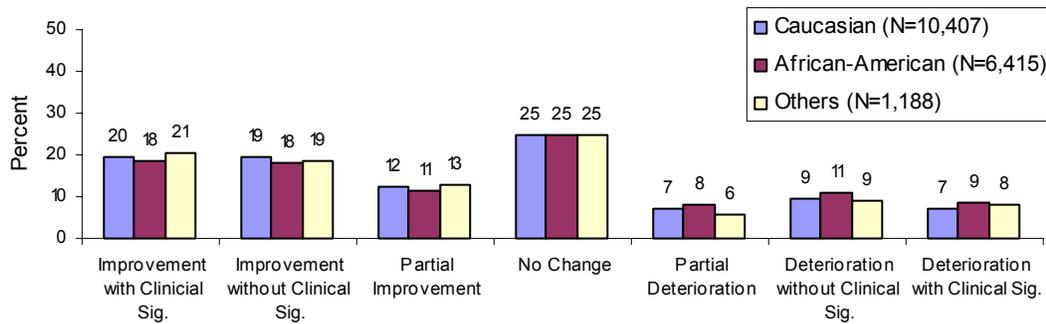
Changes by Race—Functioning

Change in Functioning by Race, Initial Assessment to 3-month Assessment

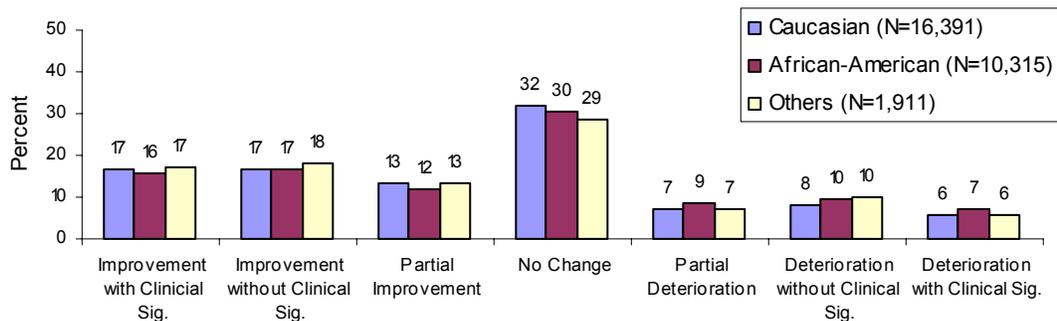
Youth Rating - Functioning (Initial - 3-month) by Race



Parent Rating - Functioning (Initial - 3-month) by Race

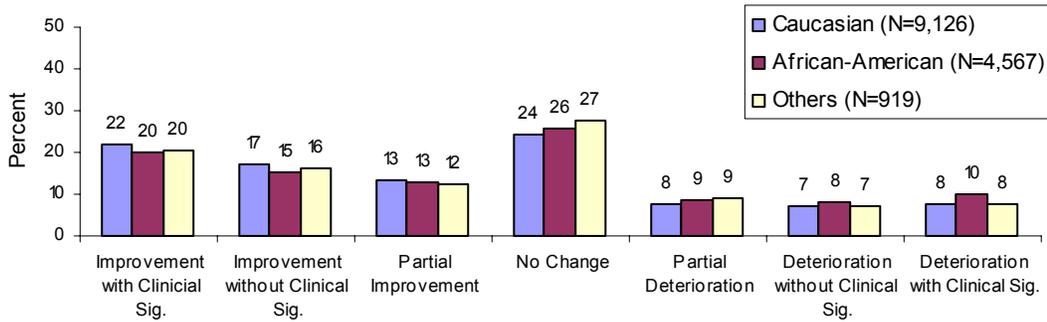


Worker Rating - Functioning (Initial - 3-month) by Race

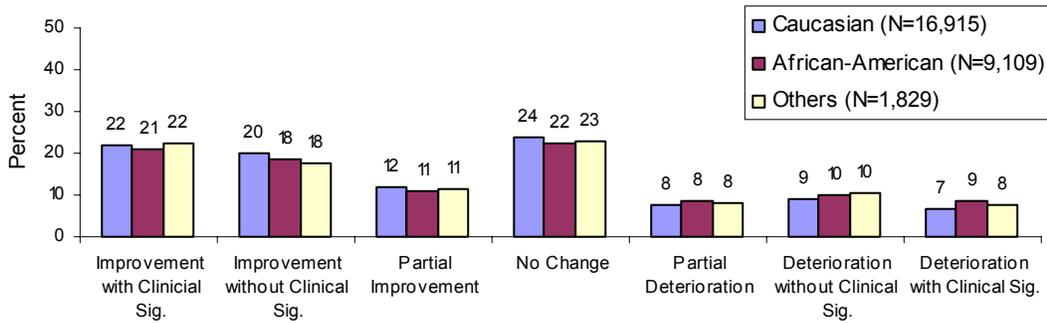


Change in Functioning by Race, Initial Assessment to 6-month Assessment

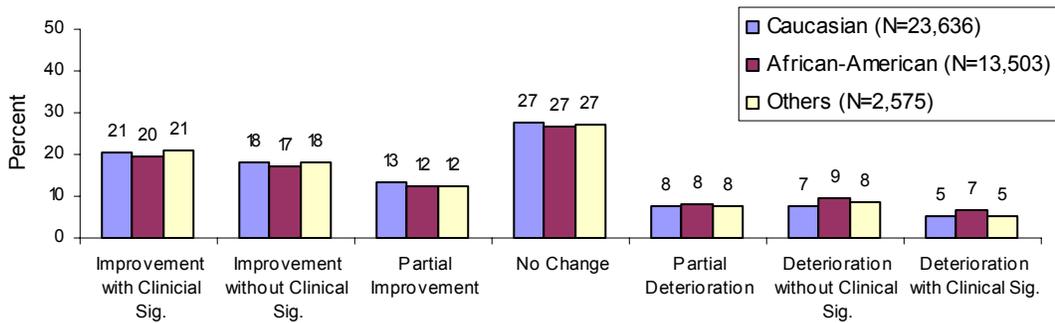
Youth Rating - Functioning (Initial - 6-month) by Race



Parent Rating - Functioning (Initial - 6-month) by Race

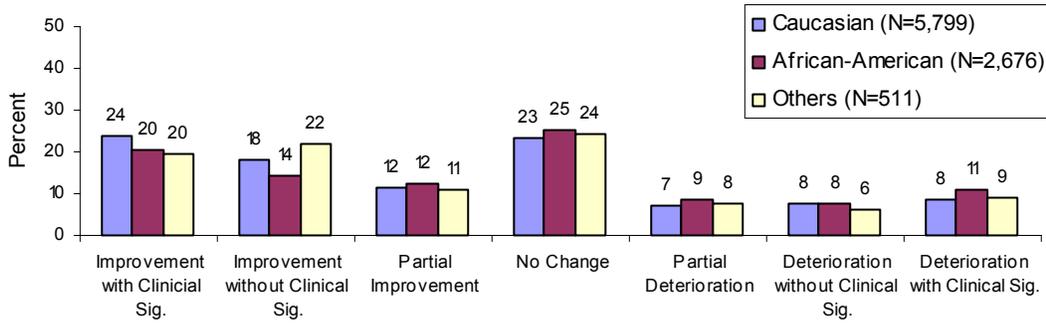


Worker Rating - Functioning (Initial - 6-month) by Race

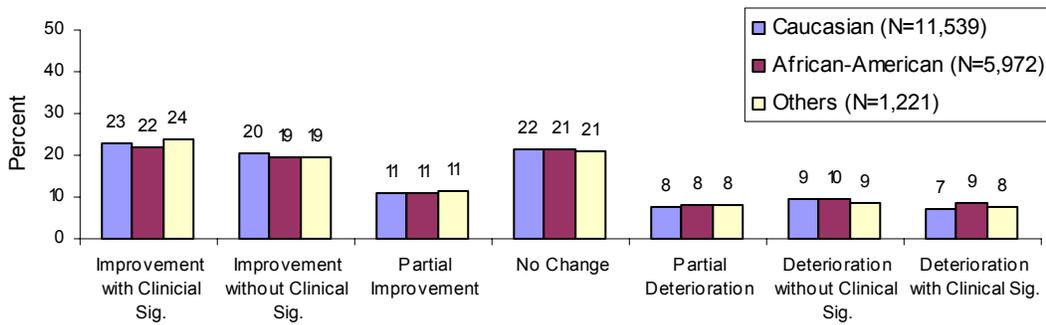


Change in Functioning by Race, Initial Assessment to 1-year Assessment

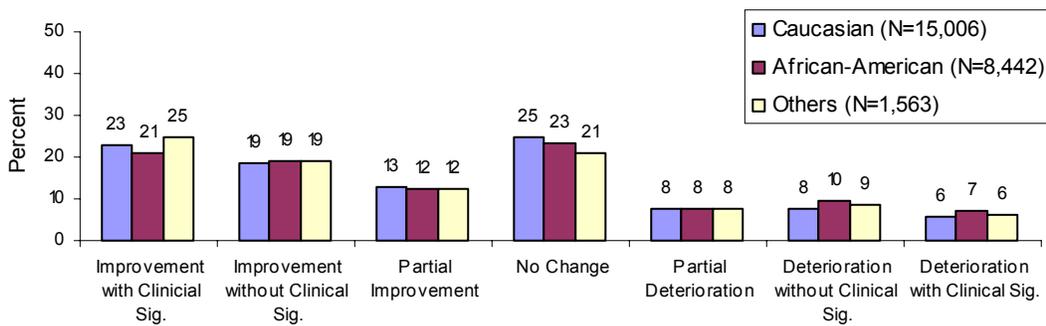
Youth Rating - Functioning (Initial - 1-year) by Race



Parent Rating - Functioning (Initial - 1-year) by Race

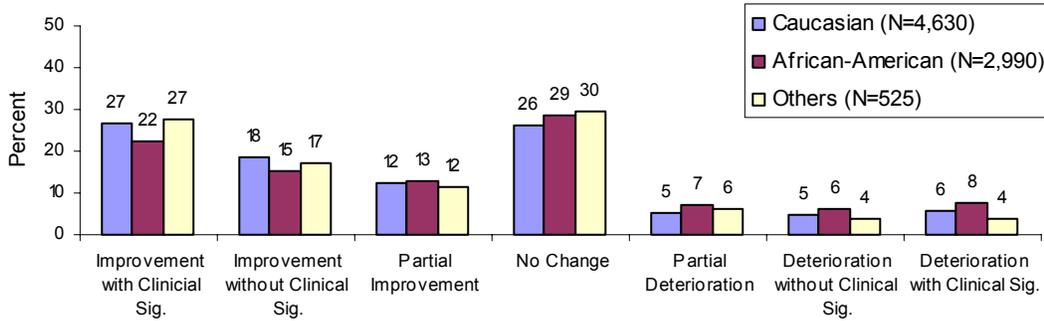


Worker Rating - Functioning (Initial - 1-year) by Race

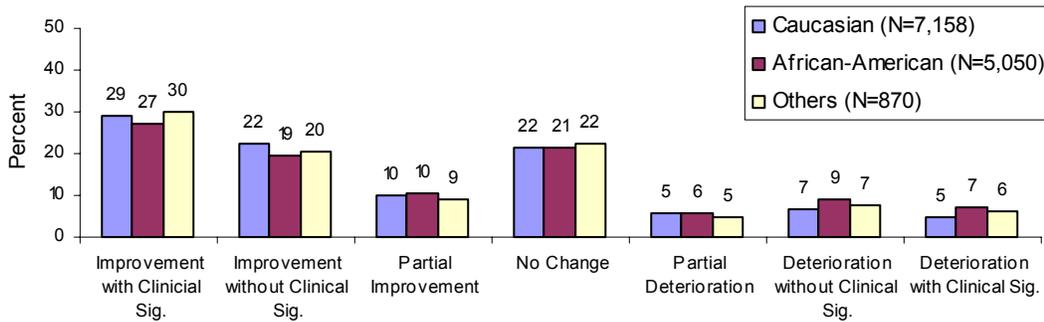


Change in Functioning by Race, Initial Assessment to Termination

Youth Rating - Functioning (Initial - Termination) by Race

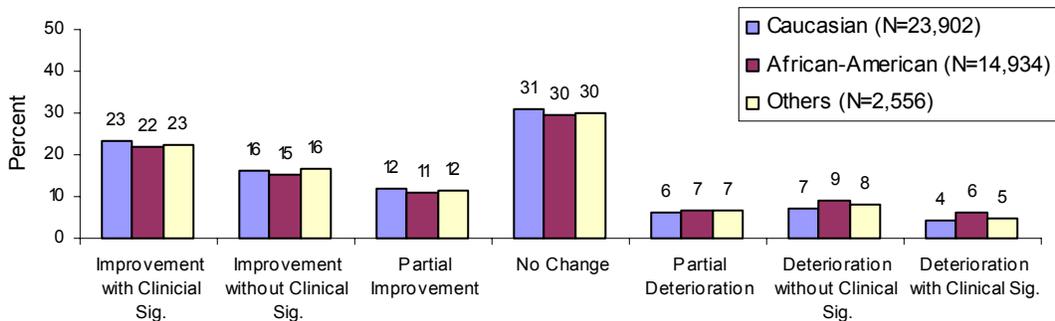


Parent Rating - Functioning (Initial - Termination) by Race



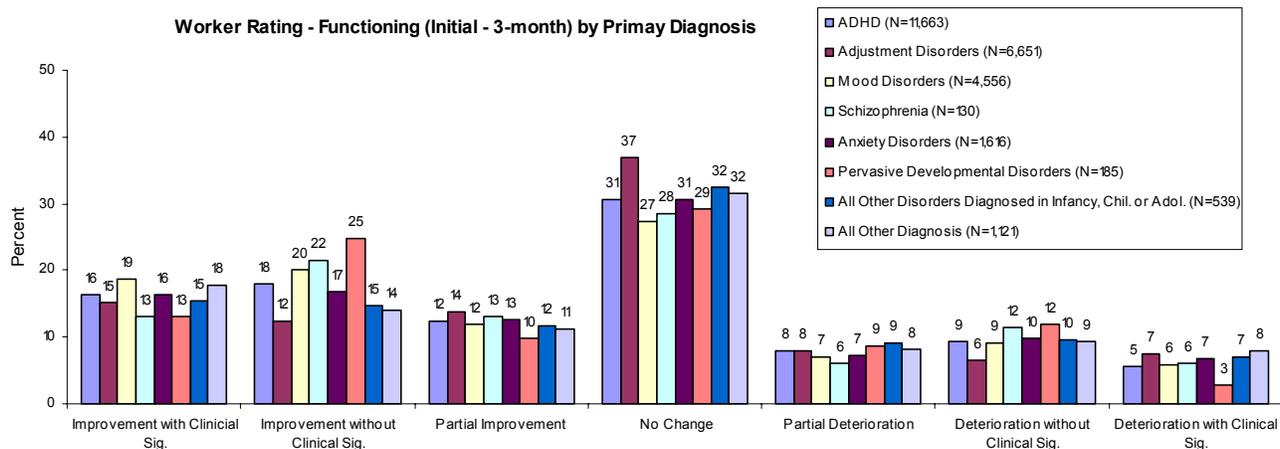
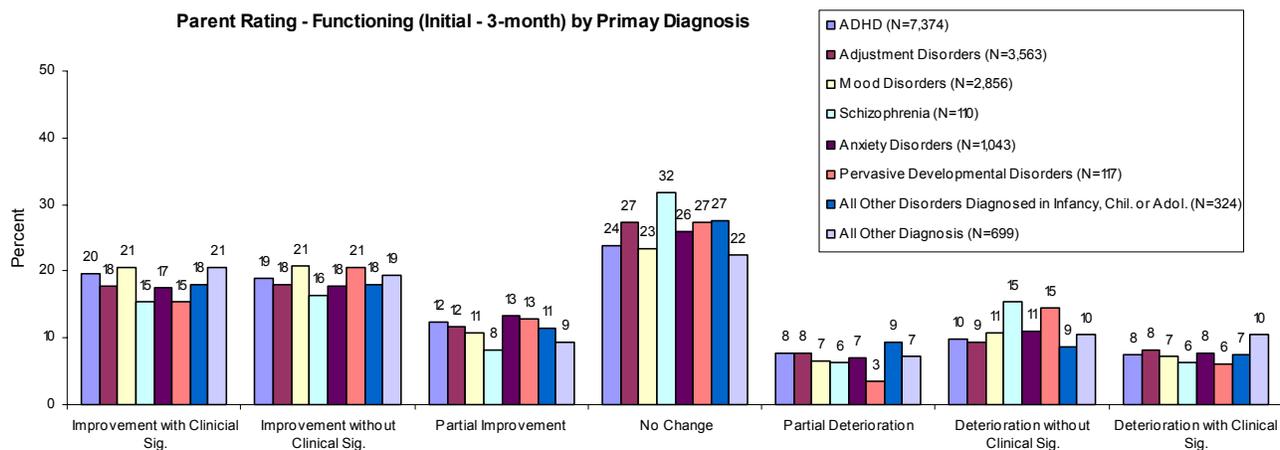
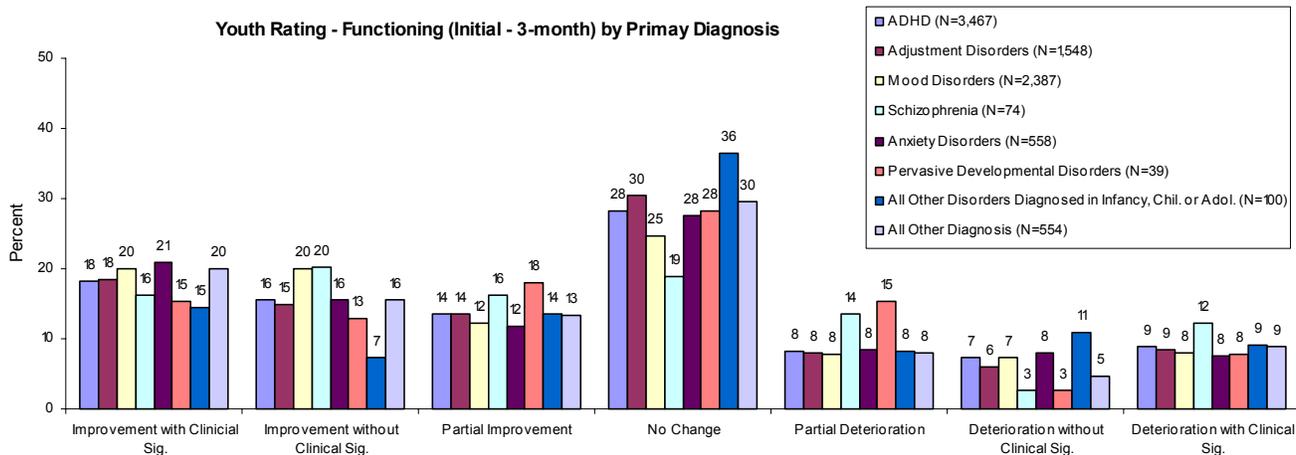
Note: Race is a significant factor in multivariate analysis with Functioning as dependent variable.

Worker Rating - Functioning (Initial - Termination) by Race

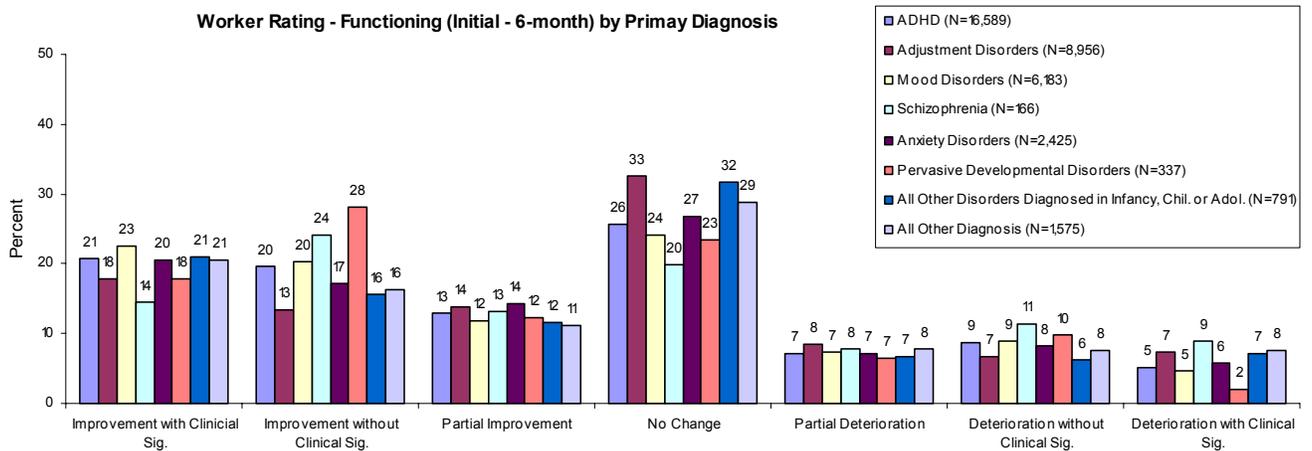
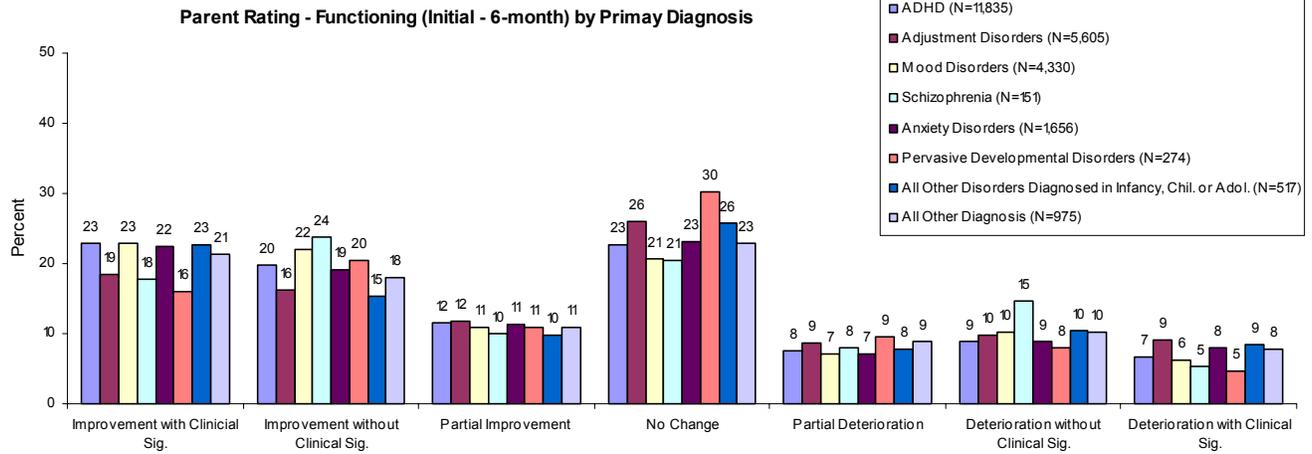
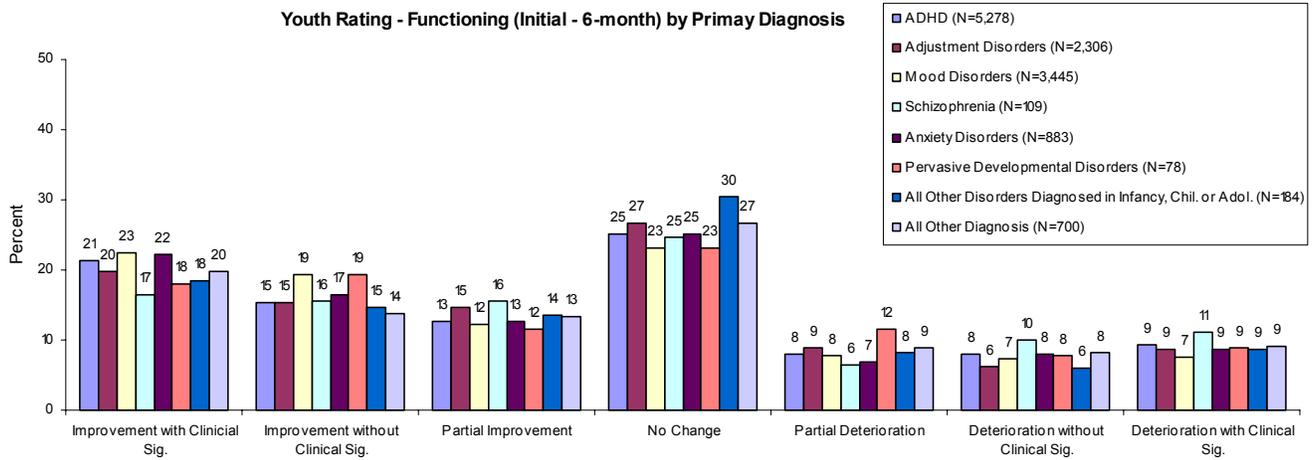


Changes by Diagnostic Group—Functioning

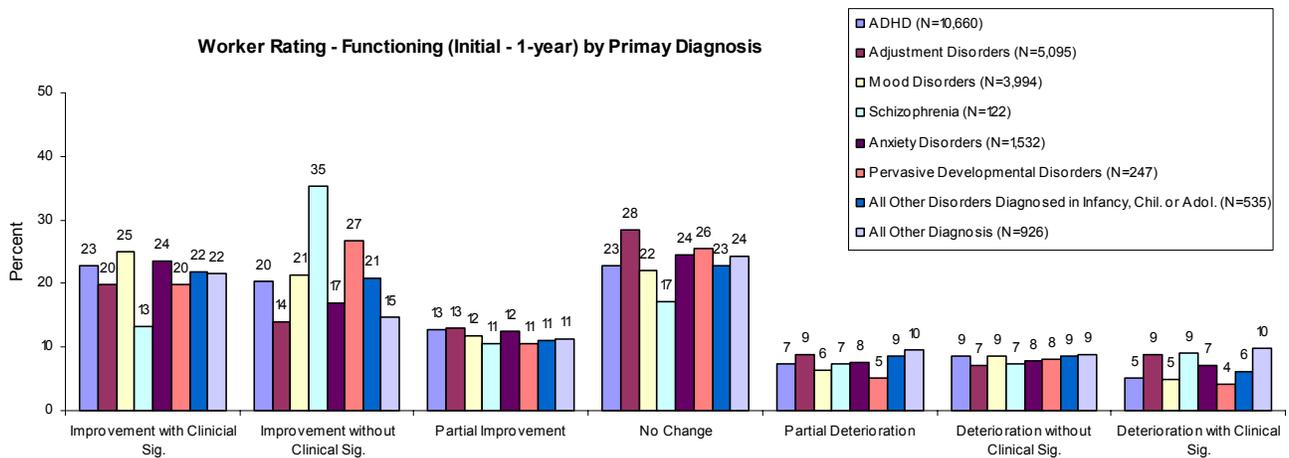
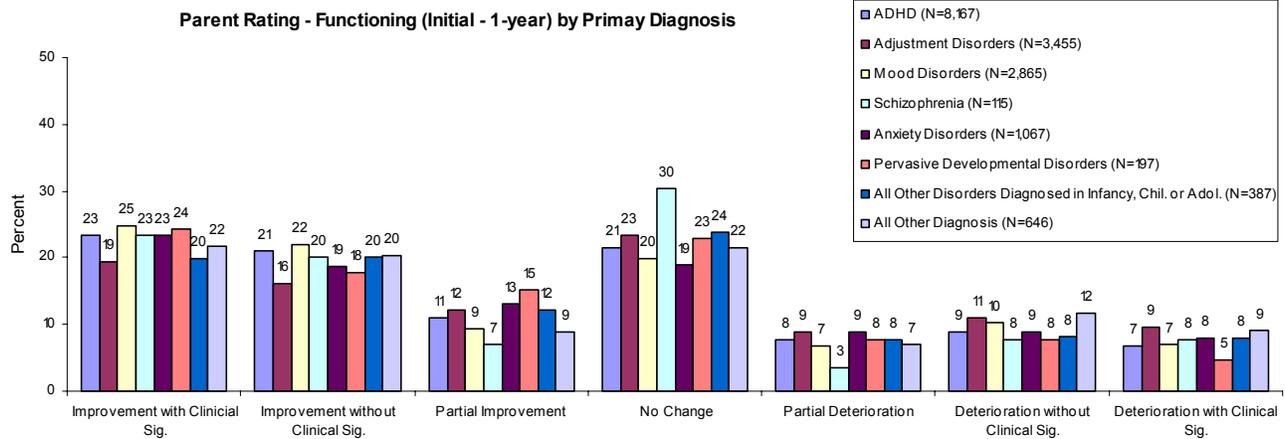
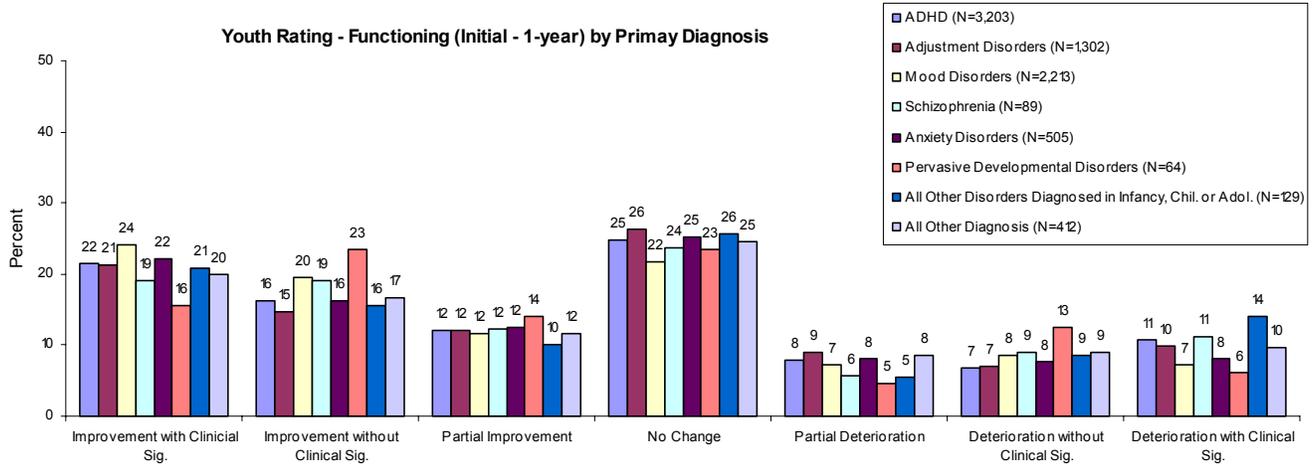
Change in Functioning by Primary Diagnosis, Initial Assessment to 3-month Assessment



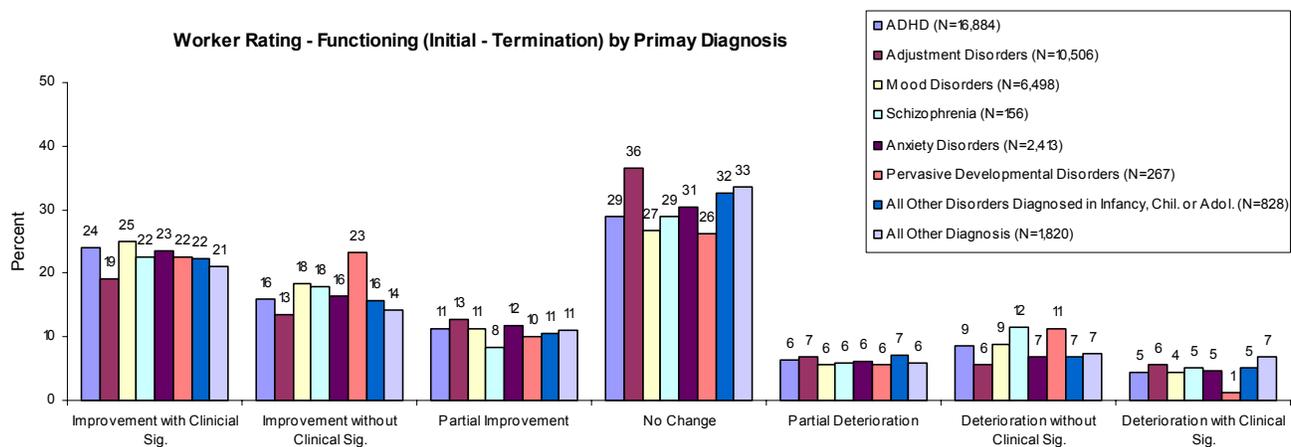
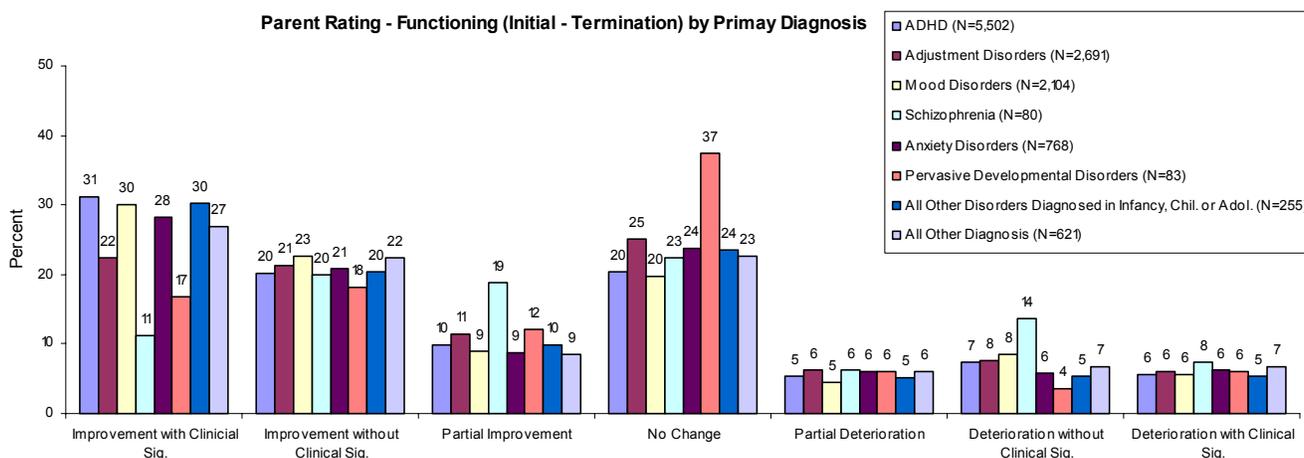
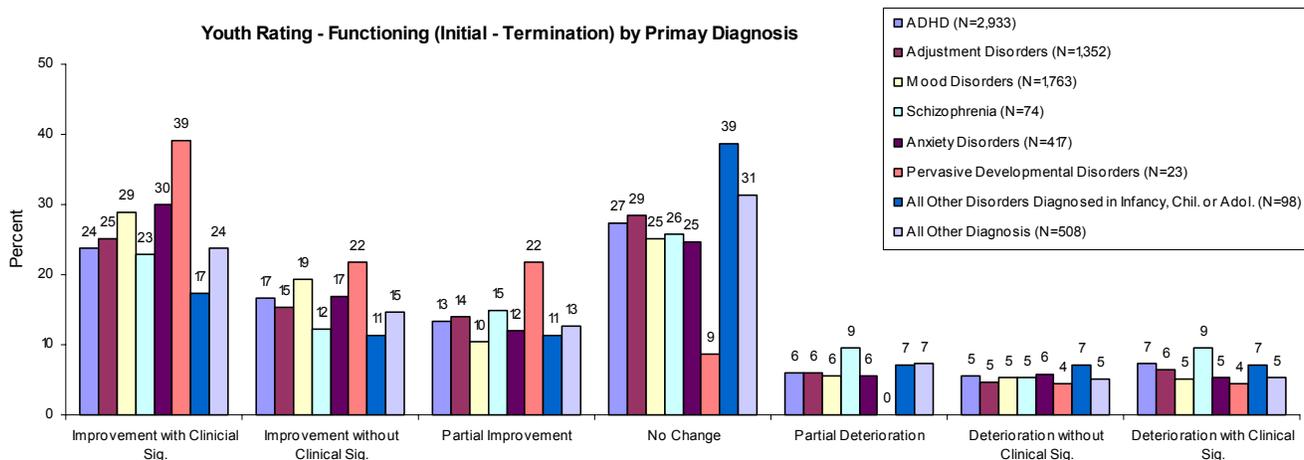
Change in Functioning by Primary Diagnosis, Initial Assessment to 6-month Assessment



Change in Functioning by Primary Diagnosis, Initial Assessment to 1-year Assessment



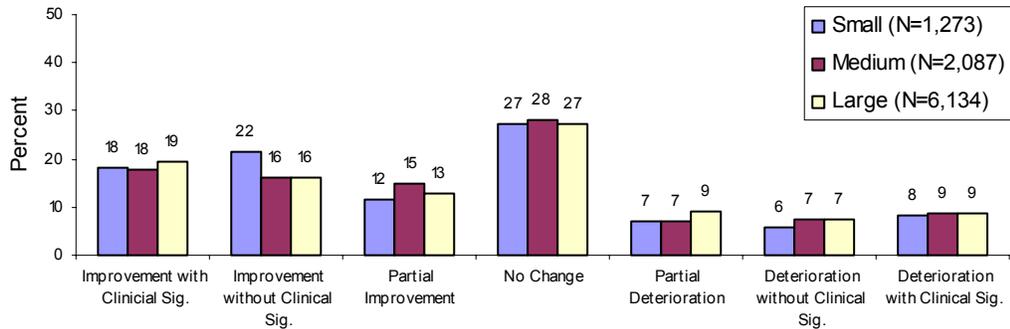
Change in Functioning by Primary Diagnosis, Initial Assessment to Termination



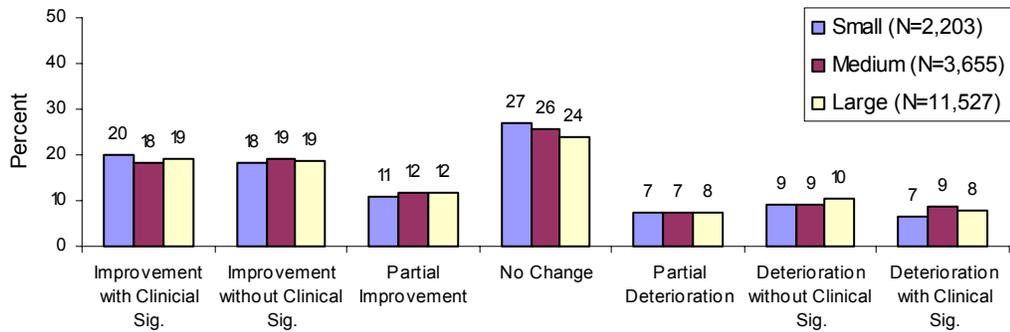
Changes by County Size—Functioning

Change in Functioning by County Size, Initial Assessment to 3-month Assessment

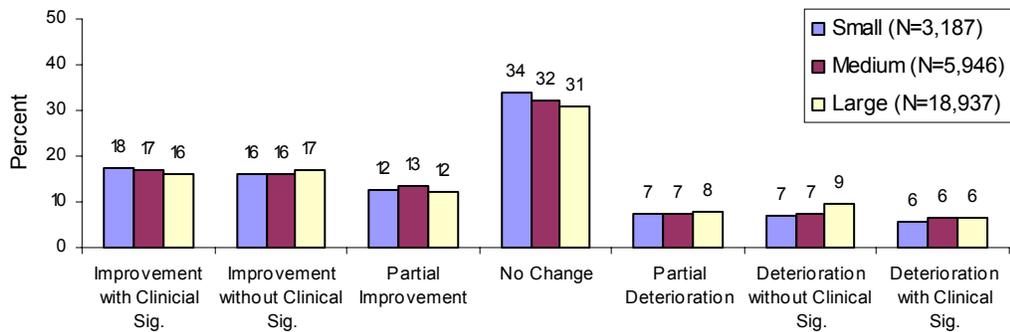
Youth Rating - Functioning (Initial - 3-month) by County Size



Parent Rating - Functioning (Initial - 3-month) by County Size



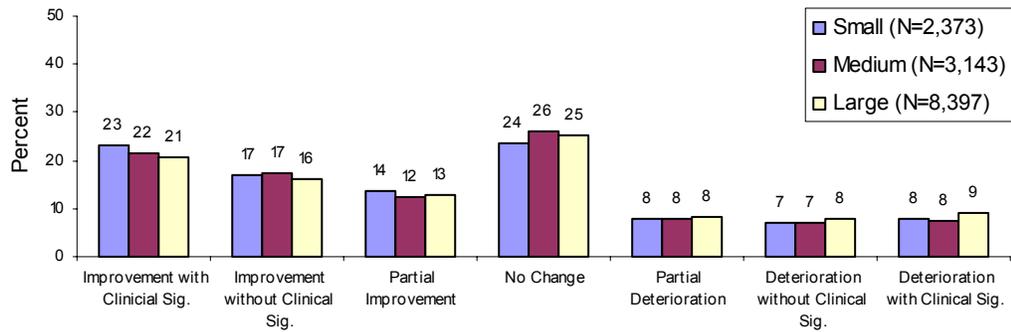
Worker Rating - Functioning (Initial - 3-month) by County Size



Note: County size is a significant covariate in multivariate analysis with Functioning as dependent variable.

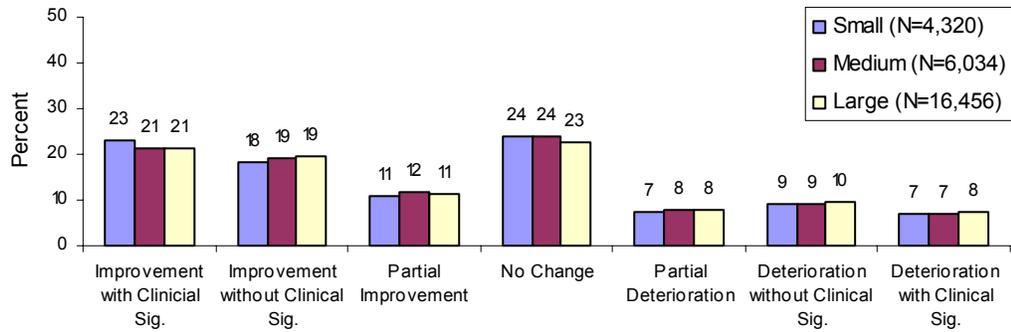
Change in Functioning by County Size, Initial Assessment to 6-month Assessment

Youth Rating - Functioning (Initial - 6-month) by County Size

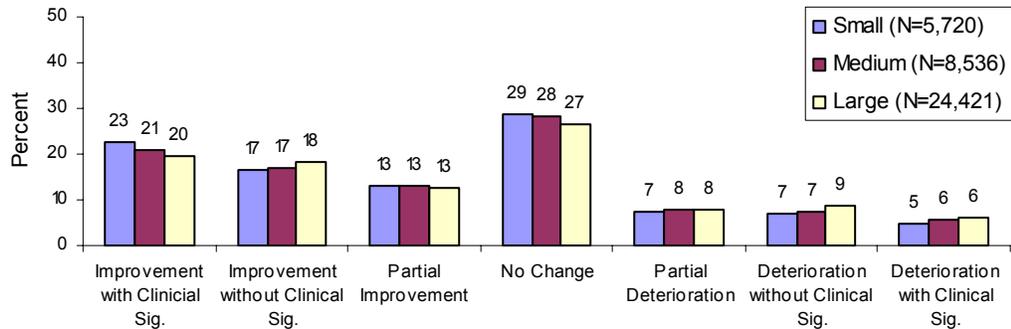


Note: County size is a significant factor in multivariate analysis with Functioning as dependent variable.

Parent Rating - Functioning (Initial - 6-month) by County Size



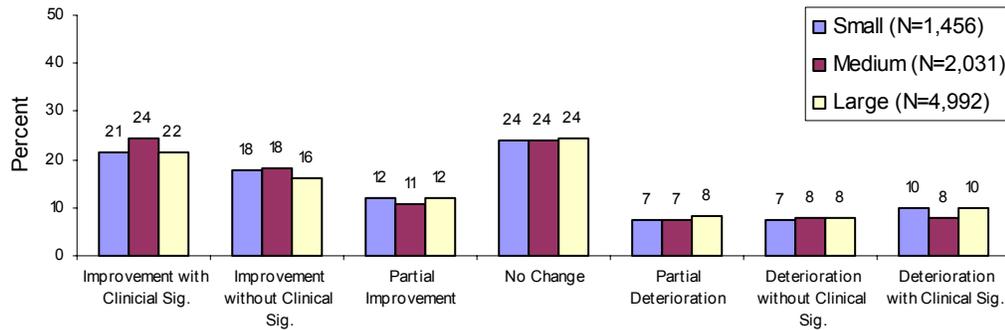
Worker Rating - Functioning (Initial - 6-month) by County Size



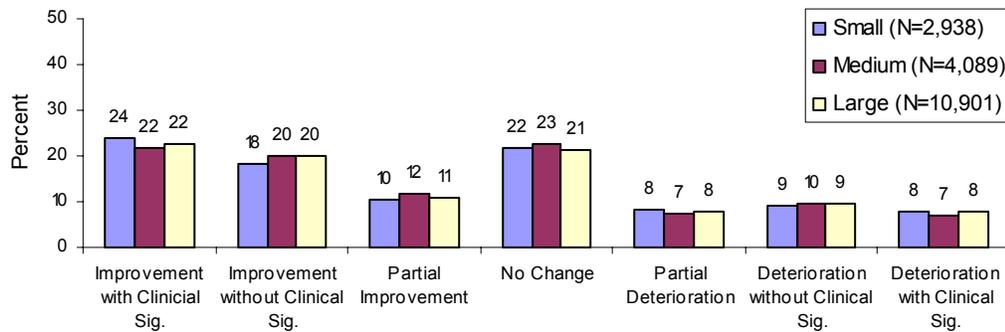
Note: County size is a significant factor in multivariate analysis with Functioning as dependent variable.

Change in Functioning by County Size, Initial Assessment to 1-year Assessment

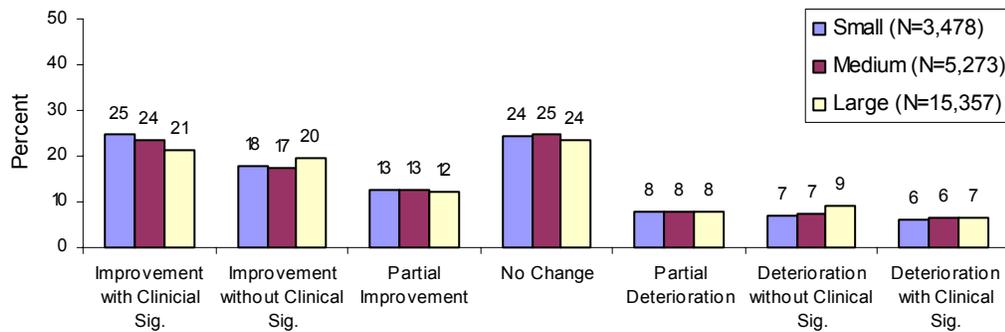
Youth Rating - Functioning (Initial - 1-year) by County Size



Parent Rating - Functioning (Initial - 1-year) by County Size



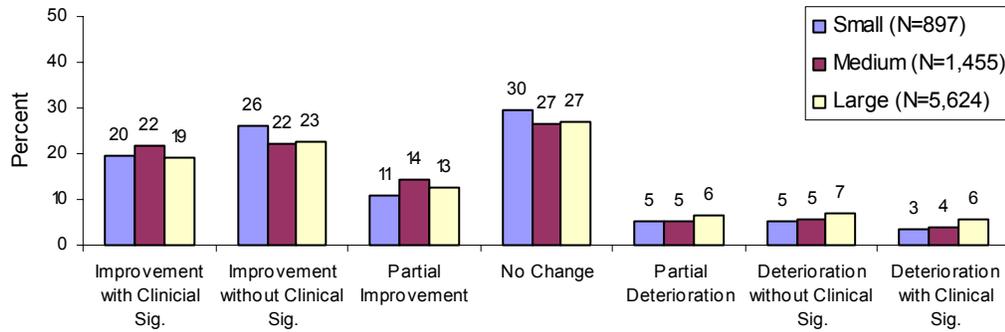
Worker Rating - Functioning (Initial - 1-year) by County Size



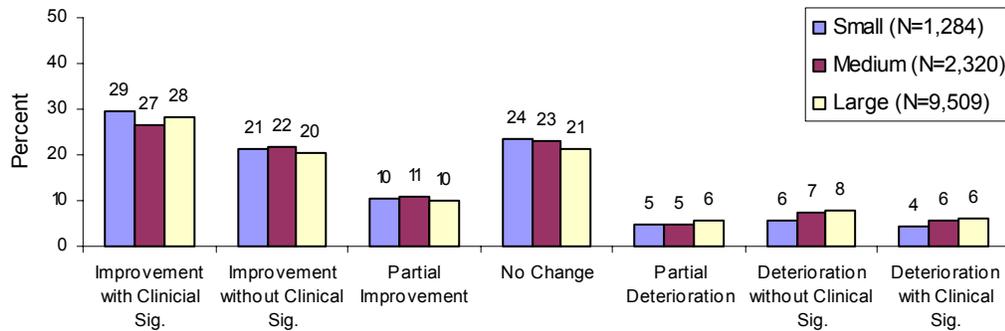
Note: County size is a significant factor in multivariate analysis with Functioning as dependent variable.

Change in Functioning by County Size, Initial Assessment to Termination

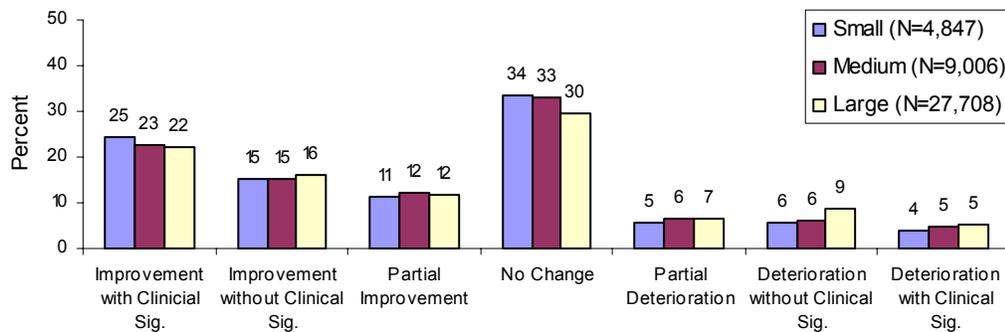
Youth Rating - Functioning (Initial - Termination) by County Size



Parent Rating - Functioning (Initial - Termination) by County Size



Woker Rating - Functioning (Initial - Termination) by County Size

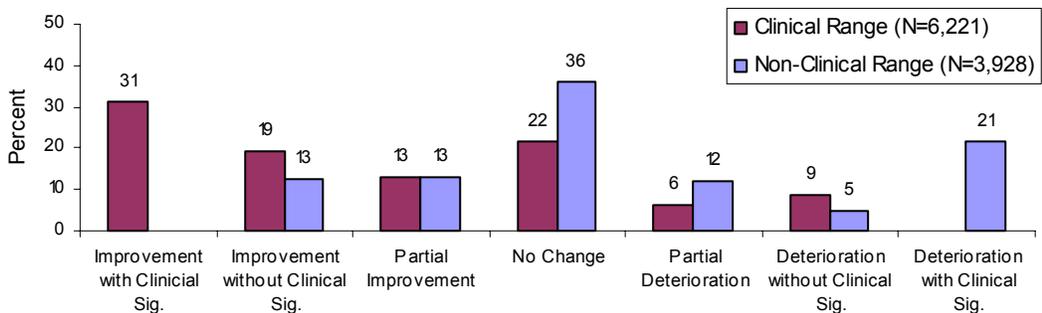


Note: County size is a significant factor in multivariate analysis with Functioning as dependent variable.

Changes by Intake Functioning Level—Functioning

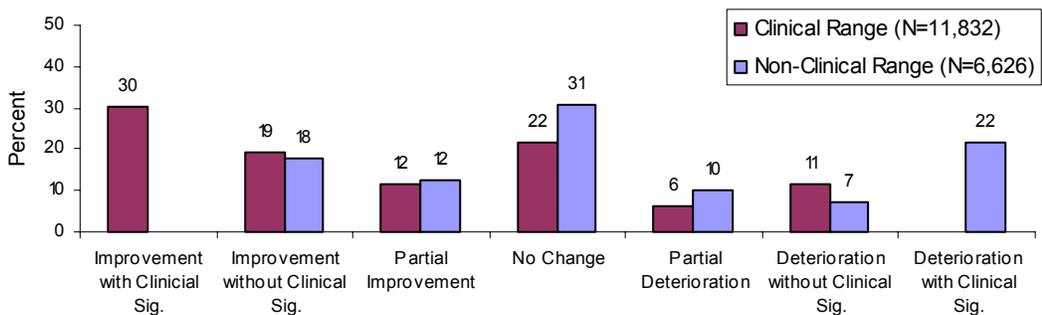
Change in Functioning by Intake Functioning Level, Initial Assessment to 3-month Assessment

Youth Rating - Functioning (Initial - 3-month) by Intake Functioning Level



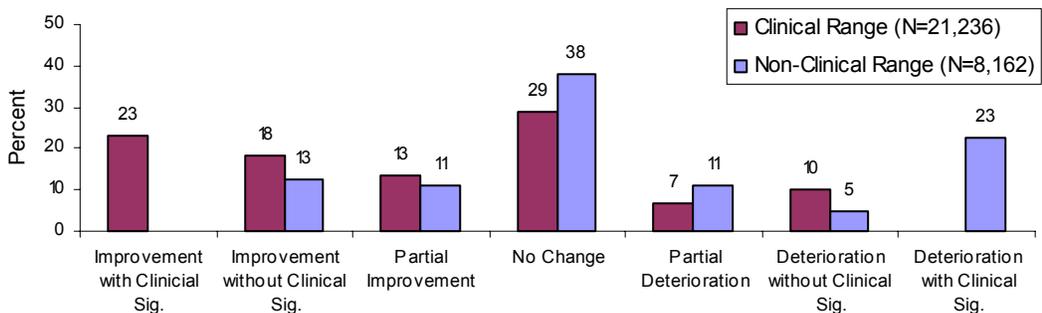
Note: Intake Functioning level is a significant factor in multivariate analysis with Functioning as dependent variable.

Parent Rating - Functioning (Initial - 3-month) by Intake Functioning Level



Note: Intake Functioning level is a significant factor in multivariate analysis with Functioning as dependent variable.

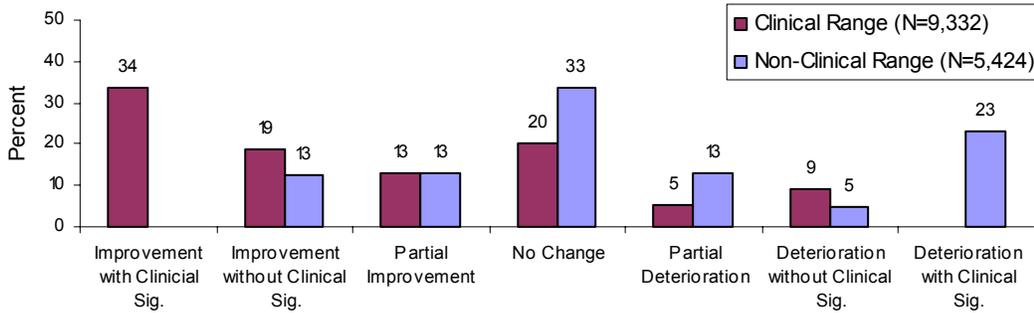
Worker Rating - Functioning (Initial - 3-month) by Intake Functioning Level



Note: Intake Functioning level is a significant factor in multivariate analysis with Functioning as dependent variable.

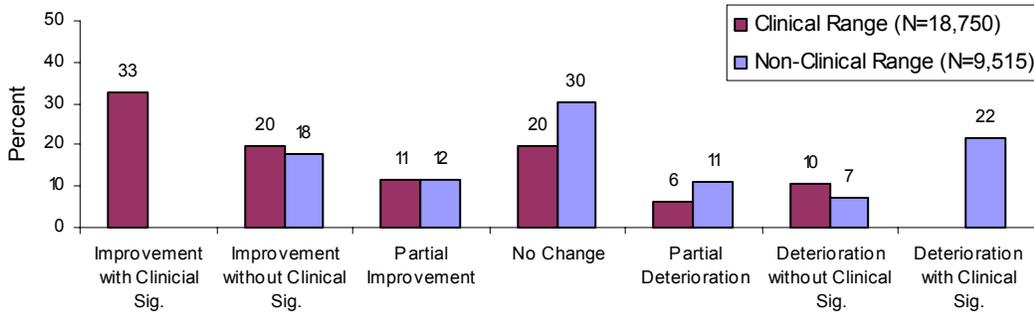
Change in Functioning by Intake Functioning Level, Initial Assessment to 6-month Assessment

Youth Rating - Functioning (Initial - 6-month) by Intake Functioning Level



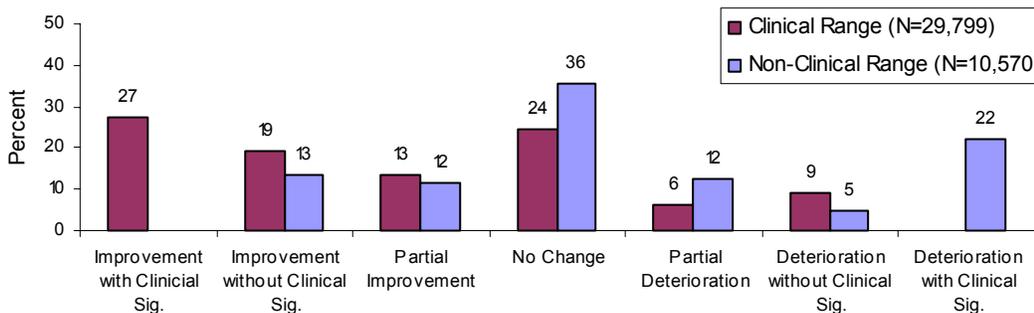
Note: Intake Functioning level is a significant factor in multivariate analysis with Functioning as dependent variable.

Parent Rating - Functioning (Initial - 6-month) by Intake Functioning Level



Note: Intake Functioning level is a significant factor in multivariate analysis with Functioning as dependent variable.

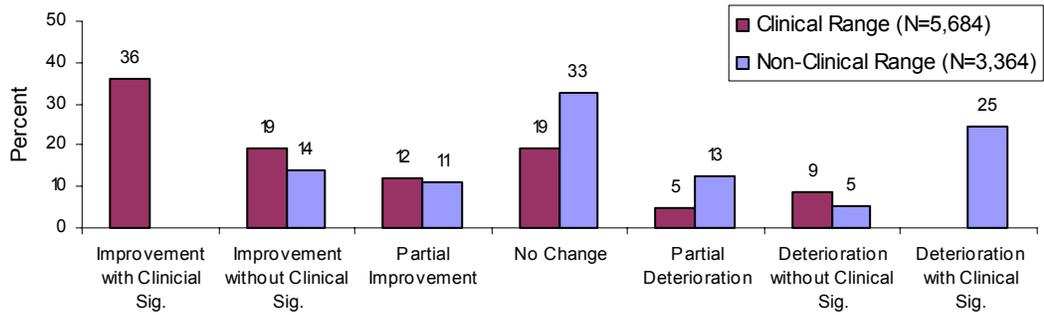
Worker Rating - Functioning (Initial - 6-month) by Intake Functioning Level



Note: Intake Functioning level is a significant factor in multivariate analysis with Functioning as dependent variable.

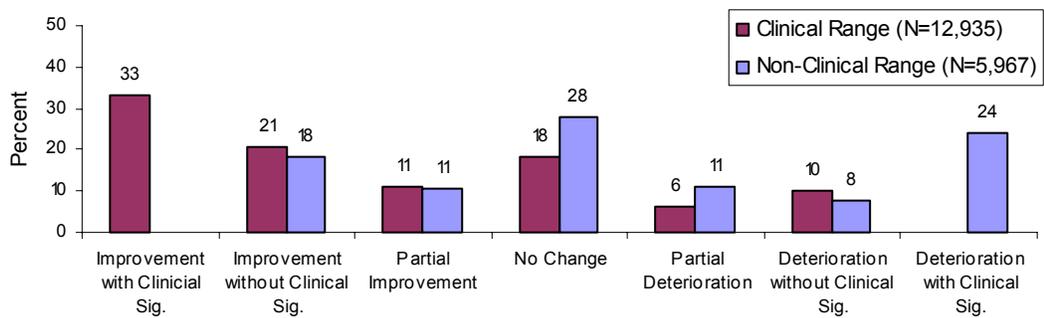
Change in Functioning by Intake Functioning Level, Initial Assessment to 1-year Assessment

Youth Rating - Functioning (Initial - 1-year) by Intake Functioning Level



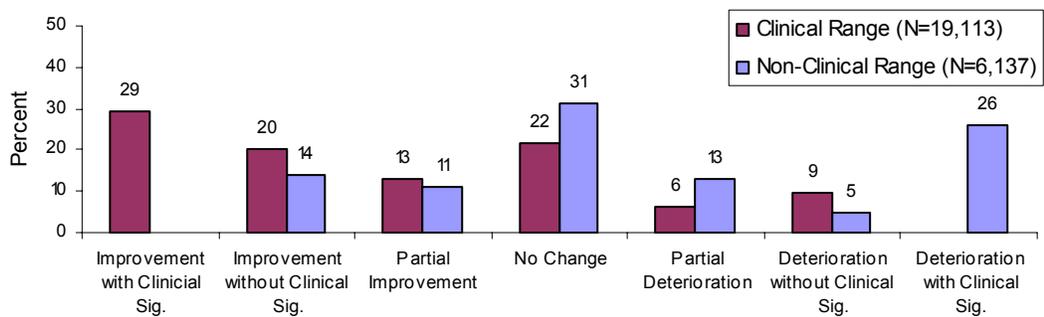
Note: Intake Functioning level is a significant factor in multivariate analysis with Functioning as dependent variable.

Parent Rating - Functioning (Initial - 1-year) by Intake Functioning Level



Note: Intake Functioning level is a significant factor in multivariate analysis with Functioning as dependent variable.

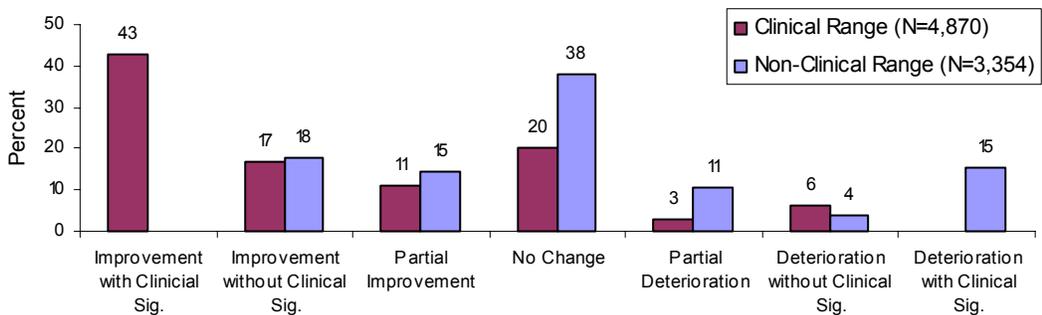
Worker Rating - Functioning (Initial - 1-year) by Intake Functioning Level



Note: Intake Functioning level is a significant factor in multivariate analysis with Functioning as dependent variable.

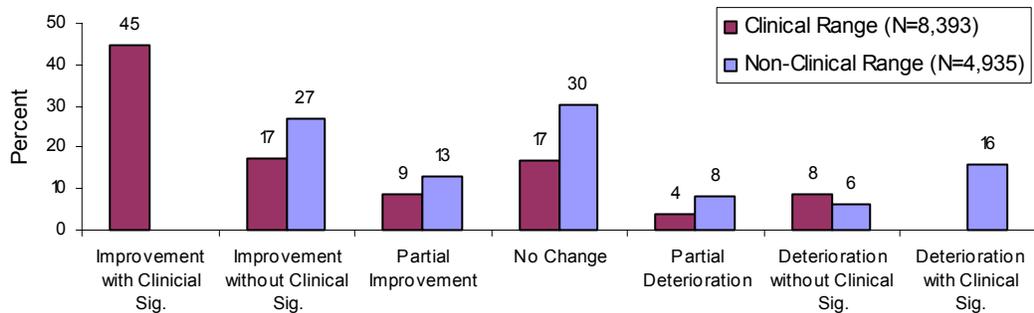
Change in Functioning by Intake Functioning Level, Initial Assessment to Termination

Youth Rating - Functioning (Initial - Termination) by Intake Functioning Level



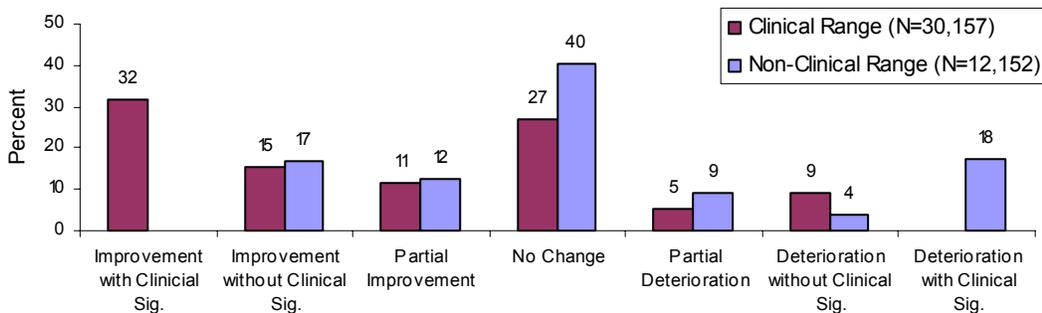
Note: Intake Functioning level is a significant factor in multivariate analysis with Functioning as dependent variable.

Parent Rating - Functioning (Initial - Termination) by Intake Functioning Level



Note: Intake Functioning level is a significant factor in multivariate analysis with Functioning as dependent variable.

Worker Rating - Functioning (Initial - Termination) by Intake Functioning Level



Note: Intake Functioning level is a significant factor in multivariate analysis with Functioning as dependent variable.

References

- Evans, C., Margison, F. & Barkham, M. (1998) The contribution of reliable and clinically significant change methods to evidence-based mental health. *Evidence Based Mental Health* 1:70-72.
- Jacobson, N. S., Follette, W. C. & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy*, 15: 336-352.
- Jacobson, N. S., Follette, W. C. & Revenstorf, D. (1986). Toward a standard definition of clinically significant change. *Behavior Therapy*, 17: 308-311.
- 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
- Lunnen, K. M., & Ogles, B. M. (1998). A multiperspective, multivariable evaluation of reliable change. *Journal of Consulting and Clinical Psychotherapy*, 66 (2): 400-410.
- McGlinchey, J. B., Atkins, D. C., & Jacobson N. S. (2002). Clinical significance methods: Which one to use and how useful are they? *Behavior Therapy*, 33: 529-550.
- Ogles, B. M., Lunnen, K. M. & Bonesteel, K. (2001). Clinical significance: history, application, and current practice. *Clinical Psychology Review*, 21 (3): 421-446.
- Ogles, B. M. & Healy, J. (2005) Calibrating the agency worker rated Ohio scales with the Child and Adolescent Functional Assessment Scale. *New Research in Mental Health*, 16: 249-255.
- Ogles, B. M., Melendez, G., Davis, D.C. & Lunnen K.M. (2000). The Ohio youth problem, functioning and satisfaction scales technical manual. Columbus, OH: Ohio Department of Mental Health, Office of Program Evaluation and Research. (<http://www.mh.state.oh.us/oper/outcomes/instruments/os.technical.manual.pdf>)
- Ogles, B.M., Melendez, G., Davis, D.C. & Lunnen K.M. (1999). The Ohio Youth Problems, Functioning and Satisfaction Scales (Short Form) User's Manual. Columbus, OH: Ohio Department of Mental Health, Office of Program Evaluation and Research. (<http://www.mh.state.oh.us/oper/outcomes/instruments/os.user.manual.pdf>)
- Wampold, B. (2001). The great psychotherapy debate. Lawrence Earlbaum Associates, Mahwah, New Jersey.
- Wise, E. (2004). Methods for analyzing psychotherapy outcomes: A review of clinical significance, reliable change, and recommendation for future directions. *Journal of Personality Assessment*, 82 (1) 50-59.

Appendix A—ANOVA Results Tables

Table 1. Significant Factors and Covariates in ANOVA Results with Changes in Problem Severity Score as Dependent Variables

	Initial to 3-month			Initial to 6-month			Initial to 1-year			Initial to Termination		
	Youth Rating	Parent Rating	Worker Rating	Youth Rating	Parent Rating	Worker Rating	Youth Rating	Parent Rating	Worker Rating	Youth Rating	Parent Rating	Worker Rating
Gender												
Age		***	***	*	***	**		**				***
Race											*	
Primary Diagnosis												
County Size		*	*	**		*						
Intake Problem Severity Level	***	***	***	***	***	***	***	***	***	***	***	***
Significant Interaction Effect	Race x Gender*; Race x Gender x Primary Diagnosis	Gender x Intake Problem Severity Level*		Race x Primary Diagnosis*; Intake Problem Severity Level x Primary Diagnosis	Intake Problem Severity Level x Primary Diagnosis**	Intake Problem Severity Level x Primary Diagnosis**			Intake Problem Severity Level x Primary Diagnosis**	Gender x Intake Problem Severity Level x Primary Diagnosis*	County Size x Gender*; County Size x Race*; County Size x Gender x Race*	
Adjusted R ²	.211	.157	.154	.232	.176	.127	.240	.173	.194	.269	.192	.157
N	8,396	15,525	25,620	12,508	24,736	35,835	7,676	16,673	22,749	6,822	11,568	38,122

* $p < .05$, ** $p < .01$, *** $p < .001$

Table 2. Significant Factors and Covariates in ANOVA results with Changes in Functioning Score as Dependent Variables

	Initial to 3-month			Initial to 6-month			Initial to 1-year			Initial to Termination		
	Youth Rating	Parent Rating	Worker Rating	Youth Rating	Parent Rating	Worker Rating	Youth Rating	Parent Rating	Worker Rating	Youth Rating	Parent Rating	Worker Rating
Gender												
Age		***	***		**	**		*			**	**
Race										**		
Primary Diagnosis												
County Size			**	*		*			*			*
Intake Functioning Level	***	***	***	***	***	***	***	***	***	***	***	***
Significant Interaction Effect		County Size x Race * ; County Size x Intake Functioning Level * ; Gender x Intake Functioning Level * ; Intake Functioning Level x Primary Diagnosis *	County Size x Primary Diagnosis ** ; County Size x Intake Functioning Level *		Race x Intake Functioning Level *	Intake Functioning x Primary Diagnosis **		County Size x Primary Diagnosis *	Intake Functioning x Primary Diagnosis *** ; County Size x Intake Functioning x Primary Diagnosis **		County Size x Race x Primary Diagnosis *	County Size x Gender x Primary Diagnosis * ; Gender x Primary Diagnosis x Intake Functioning Level *
Adjusted R ²	.168	.135	.112	.185	.148	.127	.194	.156	.142	.200	.154	.102
N	8,437	15,449	25,416	12,553	24,483	35,835	7,661	16,411	22,503	6,997	11,709	38,070

* $p < .05$, ** $p < .01$, *** $p < .001$